Validation documentation
TCM - Granulated herbal extracts (PhytoComm)

HiperScan GmbH

February 15, 2022

Validation documentation of the chemometrics database of the near-infrared spectrometer *Apo-Ident*.
Report number: **56-2.4-2022-01**
info@hiperscan.com, www.hiperscan.com

Contents

Introduction

The unequivocal proof of the identity of pharmaceutical raw materials based on a monograph or traditional alternative methods is work-intense, time-intense and economically often no longer makes sense. Near-infrared spectroscopy (NIR) provides a new means here. It enables relatively easy, fast and nonetheless safe identity testing via the preparation and evaluation of spectra.

The analysis system *Apo-Ident* was developed specially for pharmacy use. Pharmacists have the duty to check the identity of all raw materials for extemporaneous products in their pharmacies. This normally takes place based on the monographs for the respective substances in the European Pharmacopoeia. But also NIR spectroscopy is described in the European Pharmacopoeia as an identification method which, as opposed to the methods incorporated in the respective monographs, is approved for testing [1] (quotation translated)

> *with the prerequisite that the same results ("namely the determination of the identity" [2]),* (quotation translated) *are achieved as with the described methods and instruments.*

The *Apo-Ident* analysis system serves to identify raw materials for prescriptions in the defined manner according to *ApBetrO* [Pharmacies Rules and Regulations] §§ 6 and 11 at pharmacies (NIR spectroscopy as an alternative testing method). *Apo-Ident* consists of three components:

- An *NIR spectrometer*, which records the spectra of non-preprocessed raw materials in a measuring glass in diffuse reflection or transflection.
- The *QuickStep* spectroscopy software controls the instrument and records the spectra and user inputs via a pharmacy-specific software plug-in. It also generates the test protocol for documentation of the testing and storage of the printout to be signed in the pharmacy.
- The software module *IdentModule* incorporates *reference databases*. The spectra from the *QuickStep* software are presented to it for evaluation.

NIR spectroscopy is a very powerful analytical method. It is also able to establish the identity of several chemical compounds and mixes in as far as an appropriate database (technically correct: a chemometric model) was created. Identity testing with *Apo-Ident* is a very safe, very fast and easy to operate analytical method for testing a large number of raw materials.

Context of this document

The suitability of the instrument, method and database is proven as follows:

- *NIR spectroscopy as a method for identity testing:* The *Ph. Eur.* [3] describes NIR spectroscopy in *Section 2.2.40* as an analytical method which is also suitable for the identification of raw materials. Therefore, validation of the method as such is not necessary.
- *Performance of the instrument:* The *Ph. Eur.* [3] furthermore describes the apparatus and the testing of its performance in *Section 2.2.40*. The document *Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident* [4] compares the implementation by *Apo-Ident* with this monograph in order to prove that *Apo-Ident* meets the specifications of the Pharmacopoeia. Each individual instrument delivered to a pharmacy is qualified in accordance with the tests described in *"Control of Instrument Performance"*. In this test, the unit consisting of analysis instrument hardware and the *QuickStep* spectroscopy software is assessed. The result is documented in a test protocol which is kept at the pharmacy.
- *Validation of the database* is documented separately for each substance class. The report at hand documents the substance class *TCM - Granulated herbal extracts (PhytoComm)*.

The *Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD)* [Working Group of German Pharmacy Inspectors] has clarified the following in its resolution dated October 16, 2013 ([5], quotation translated):

> *NIR is a testing method incorporated in the Pharmacopoeia. The testing quality depends on the quality of the database stored. The* APD *views the use of NIR instruments in case of ensured validation of the databases used in conjunction with it as one of several options for identity testing.*

The *APD* defined more precisely ([6], quotation translated) on October 1, 2014:

> *The use of near-infrared is a recognised testing method according to* Ph. Eur. 8. *For the use of NIR instruments in pharmacies for testing the identity of raw materials, sufficient and verifiable validation of the instrument used is required. The quality of the database stored by the instrument manufacturer is decisive for quality. Batch-specific differences with the same original substances must be taken into account if present.*

So NIR is basically suitable. The validity of the reference database is proven with the existing validation documentation.

Criteria for the inclusion of substances

This validation documentation describes the results of the validation of the reference database for the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Validation documentation is created for each published version of the reference database for all substance classes incorporated.

The reference database is incorporated in the software module *IdentModule*. During identification testing with *Apo-Ident*, spectra which are used for evaluation purposes are presented to it by the *QuickStep* software. In the same manner, the *IdentModule* is presented all validation spectra successively during the validation runs for evaluation purposes. The *IdentModule* responds respectively (without taking the initial assumption into account) with the identified substance or rejects it as unknown. The correctness of this response is checked for each possible initial assumption and counted.

The results are summarised for each substance and reproduced in this document. The core statement of this validation report is that the following criteria must be fulfilled for each database entry, so that *Apo-Ident* offers verification of identity by means of NIR for the relevant substance/substance group:

- The database is exclusively generated from spectra which have been recorded by *HiperScan GmbH* on traceable samples in pharmaceutical quality.
  - The samples are procured via typical pharmacy sources (*DAC III.2.: Bezugsquellennachweis für Rezepturbestandteile [reference source for prescription components]* [7]).
  - A valid manufacturer's certificate exists (content, purity and identity of the batch).
  - The identity was confirmed by a certified test laboratory or *HiperScan GmbH*.
- Each version of the reference database (every update) is validated in-full.
  - Calibration spectra (*Type A*), other spectra recorded under the control of *HiperScan GmbH* (*Type B*), and spectra from the field (*Type C*) are presented to the *IdentModule* for evaluation in three separately evaluated validation runs.
  - Here, no single *false positive* result may arise.
  - Here, the various substance classes are also tested for reciprocal rejection, where this is objectively justified (see *Summary* section).
- In the validation with spectra recorded under the control of *HiperScan GmbH*, spectra of at least one independent sample must be considered, i.e. spectra from a batch whose spectra have not been used for the generation of the database. In addition, the set of *Type A* and *Type B* spectra must originate from at least three different batches.
- Spectra of additional substances may be used for the generation of the database even though they will not be offered for identification by this database. The purpose is the reliable distinction from these substances.

- Any positive result of *Apo-Ident* confirms the identity of the substance/substance group and distinguishes it from all other substances of the database. In the case of substance groups, the result is ambiguous: Distinction from all substances not belonging to the group is proven. The substance is identified as a member of this group. However, within the substance group, it is not possible to reliably classify which substance has been tested.

- The criteria for clear identifiability are a specificity of 100 % (true negative rate) and a minimum distance in the distance matrix. See 2. d) under Model creation procedure and validation runs.

Validation concept

Chemometrics is a statistical technique for the extraction of relevant chemical information from spectra. In mathematics, this method is described as *multivariate data analysis*. Chemometrics proceeds here as follows:

1. Collection of spectra for the *calibration sample*. The results (identities) of the calibration sample must be known. The calibration samples must be representative for the samples which are to be evaluated later. Therefore, they must take the various possible (physical) compositions into account. (Therefore, sourcing calibration samples for NIR from the specialist trade is superior to the use of CRS reference substances.)

2. The first mathematical step is *calibration*. Here, the chemometric model is calculated from the *calibration sample spectra* (reference spectra) and limits as well as some parameters are stipulated. The chemometric model is used later to calculate the analysis result (*prediction*).

3. Collection of further spectra for the *validation sample* which should be independent of the *calibration sample*. The results (identities) of the *validation sample* must also be known. The textbook suggests a random sample with a normal scope of 25 % to 50 % of the *calibration sample* [8].

4. The second technical data step is *validation*. Here, the chemometric model created is evaluated based on the spectra of the *validation samples*. As validation parameters for the identification, the *Ph. Eur. Section 2.2.40* [3] specifies the specificity and robustness.

The validation step according to the textbook has the target of estimating the performance capability of the model created based on a random sample. In order to achieve the best possible precision, attention is paid to the calibration sample. In the field of pharmaceuticals, the safety of the method has priority. In order to be able to *validate* the model within the regulatory scope, the validation step must include probative force. For this purpose, the validation sample must be *representative and complete* in order to enable the testing of all cases.

A *sufficient number of batches* must be secured for validation because validation finally proves whether the number of batches in calibration suffices.

Each substance is validated individually. The validation results are documented per substance in this document. Moreover, the documents show how many and which batches have been used for creating the model or model validation.

At least one certificate is taken in for each substance from an accredited test laboratory for the independent testing of identity of the sample. The identification number of the corresponding test certificate is listed in the report, enabling traceability of a substance tested according to the monographs in the Pharmacopoeia.

Model creation procedure and validation runs

The safety of the chemometric models is guaranteed by several measures during model creation, of which the validation step is the final one. Normally, the procedure is as follows. It is in particular valid for the active pharmaceutical ingredients (APIs) *solid API excipients*, *liquid/semi-solid API excipients (with a test certificate)*, *narcotics - solid medicinal substances* and *drugs*. If, for individual substance classes, variations are required, they are depicted in the section *Particularities of individual substance classes*.

1. Collecting the reference spectra (calibration sample)

a) Procurement of the samples from the same sources from which pharmacies source their raw material for compounding (Caelo, Fagron, Euro-OTC, . . . , see also *DAC III.2. Bezugsquellennachweis für Rezepturbestandteile* [Sources of supply for compounding] [7]).

b) Testing the suitability according to *ApBetrO* [Pharmacies Rules and Regulations] §§ 6, 11, that is to say the availability of a valid manufacturer certificates via identity, purity and contents of the batch.

c) Recording standard 40 spectra of the sample in different positions, as a standard on four instruments. Here, handling and presentation of the samples as later in the pharmacy.

d) Visual checking for anomalies in the spectra. In case of indications of measurement errors, measurement must be repeated. If a signature is missing in the spectrum, the substance may be excluded from the start as not promising (the spectra are nonetheless entered in the database validation as independent *Type B* spectra).

e) Testing identity. For each substance, a certificate of correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the respective following substance page of this validation documentation the *Mahalanobis distance* to this reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance. Such samples underline the statistical spread of the original reference substance, but cannot add any new characteristics of the substance.

   HiperScan GmbH cooperates with some suppliers as follows: the raw materials supplier takes a sufficiently large sample in his incoming goods area so that part of it can be used for recording the NIR spectra. The remainder of the sample goes to analytics for market approval. The manufacturer's batch certificate emerges from these identity, contents and purity tests, which consequently also prove the correct identity of the NIR reference sample. Therefore, the NIR spectra are suitable for structuring the database *(Type A)* and can also optionally be used for validation *(Type B)*. The samples which this applies to are marked in the validation report with a footnote.

f) If the identity of the new sample is proven, it is declared as a reference sample and the spectra are approved for structuring the database.

2. Generating the chemometric models (calibration)

a) Determination of the transformation matrix from the reference spectra using variance maximisation [8, 9]. (*All* reference spectra are always included, even if only a few spectra are added for an update.) All reference spectra receive the same data pre-treatment, which is also later applied to all measurement spectra in the field (in the pharmacies).

b) Checking that the number of principal components used is still sufficient.

c) Calculating the limits for each substance from the spread of the reference spectra. The calculation regulation is identical for each substance in a substance class.

d) Überprüfen der Abstände zwischen den Grenzen der trennbaren Substanzen: Die Distanzmatrix enthält die *Mahalanobis-Abstände* von jeder Substanz zu jeder anderen. Die Werte hin und zurück sind jeweils unterschiedlich, weil die Streuung der Ausgangssubstanz eingeht. Ist eine Distanz kleiner als der Mindestabstand, so gelten die Substanzen als nicht sicher trennbar. Der Mindestabstand ist auf 9 festgelegt. Der Entwickler des Modells darf einen größeren Mindestabstand festlegen (ein Wert für das gesamte *chemometrische Modell*), um die Trennschärfe zu erhöhen.

e) Überprüfung des Modells anhand der Referenzspektren. Es sind keine *falsch-positiven* Ergebnisse erlaubt.

f) Wird eines der Kriterien verletzt (d) *Unterschrittener Mindestabstand zwischen zwei Substanzen* oder (e) *Eine Substanz wird als eine andere identifiziert*, entscheidet der Entwickler der Datenbank, welche der folgenden Optionen er anwendet:

- Er nimmt beide Substanzen aus der Datenbank. (Die Spektren bleiben in der Validierung und dürfen auch in den Aufbau eingehen. Sie werden aber nicht zur Prüfung angeboten.)
- Er bildet eine Substanzgruppe mehrerer nicht sicher trennbarer Substanzen. Dann ist das Ergebnis mehrdeutig: Das chemometrische Modell stellt fest, dass es sich bei der Probe um eine der Substanzen aus der Gruppe handelt und dass es sich um keine andere Substanz handelt. Es kann aber nicht sagen, um welche der Substanzen es sich handelt. Um die eindeutige Identität festzustellen, muss der Anwender eine geeignete ergänzende Prüfung durchführen.
- Er erstellt ein weiteres *chemometrisches Modell* mit geringerem Umfang, in das mindestens alle Substanzen der nicht sicher trennbaren Substanzgruppe eingehen (Zweite-Stufe-Modell). Zweite-Stufe-Modelle werden nur aufgerufen, wenn die erste Stufe festgestellt hat, dass es sich nur um eine der Substanzen handeln kann, die in den Aufbau der Zweiten Stufe eingegangen ist.

3. Set of validation spectra (validation samples)

The following is provided for validation:

a) *Type A*: The reference spectra = calibration spectra from which the database was generated. These also include spectra from substances which the *chemometric model* should not identify, but were also recorded during generation in order to increase selectivity. (As a result, the model "learns" to differentiate from other substances which are actually unknown to it.)

b) *Type B*: Spectra recorded under the control of *HiperScan GmbH* and not used for the generation of the database. These also include reference spectra of other substance classes, and spectra that are not used as reference spectra. Samples are considered to be independent, if they originate from a batch, of which no spectra have been used for the generation of the database. (Up to *IdentModule 2018-01*, samples were still considered independent if the sampling was done independently, i.e. if they originate from another sales container.)

c) *Type C*: Spectra from the field, which have not been recorded under the control of *HiperScan GmbH* and have not been used for the generation of the database. The spectra include not only substances of the substance class to be tested, but also substances from other classes.

All manufacturers' batches from which spectra are used for the validation are listed by substance in this document: for substances included in the substance class *TCM - Granulated herbal extracts (PhytoComm)* in the respective validation reports; otherwise in attachments *A*, *B* and *C*.

Furthermore remains valid: validation spectra may only be removed if a error in the spectrum can be proven. Here, the spectra are not deleted, but instead placed on a *blacklist* incorporating the reason, date and initials in the commentary.

The section *Particularities of individual substance classes* treats the other substance classes from which *Type B* and *Type C* spectra are cited for validation purposes.

4. Validation runs and approval

a) Validation spectra are transferred holistically to the *IdentModule* for evaluation in the same way as the spectroscopy software *QuickStep* transfers measured spectra.

b) Following the provision of each spectrum, the *IdentModule* responds as to whether it has recognised a substance and which substance was recognised.

c) The correctness of this response is checked for each possible initial assumption (each measurable substance with the substance class) and counted according to *true negative*, *false negative*, *true positive* and *false positive*. These figures are provided for each substance and additionally in the section *Summary*, separated according to types *A*, *B* and *C* .

d) No *false positive* results whatsoever are permissible.

e) If the criterion is also met for all substance classes, the *IdentModule* is approved.

Particularities of individual substance classes

Basically, *HiperScan GmbH* procures and tests the manufacturer's certificate for the batch, commissions external testing of the identity of the sample or carries it out independently and stores the certificates. As described, this process is established for the Pharmacopoeia substances, that is to say for substance classes **APIs & excipients, solid**, **APIs & excipients, liquid/semi-solid (with analysis certificate)**, **Narcotic substances, liquid/semi-solid** and **Drugs**. Therefore, *HiperScan GmbH* is able to furnish proof of the identity of the reference samples. In case of manufacturer-specific substance classes and others, individual steps are organised differently in-part:

The substance class **APIs & excipients, liquid/semi-solid (other)** (often described as cosmetics) incorporates substances for which no specification of the requirements of the pharmaceutical quality is determined, neither in a Pharmacopoeia monograph, a DAC/NRF monograph nor via a manufacturer's specification. Consequently, neither the identity nor contents can be tested independently. No certificates whatsoever exist for the reference samples. So here, merely the matching of the sample with former samples of this product is established and confusion with the other substances is ruled-out. (If the manufacturer of such a substance prepares a specification, determines testing methods and provides manufacturer's certificates in accordance with *ApBetrO* [Pharmacies Rules and Regulations] §§ 6, 11, *HiperScan GmbH* can assign the substance to the substance class *APIs & excipients, liquid/semi-solid (with analysis certificate)* again in the future).

Substance class **HCK – nutritional supplements (Hepart)** contains the HCK micro-nutrients from the Swiss company *Hepart AG*. *HiperScan GmbH* receives the reference samples directly from the manufacturer. For each reference sample, *HiperScan GmbH* also receives manufacturer's certificates and keeps these. New checking of the identity of the reference sample is not carried out by *HiperScan GmbH*. The identity of the reference samples is therefore documented by *Hepart AG*. The spectra of all batches provided by *Hepart AG* are recorded by *HiperScan GmbH* and entered in the database.

All the manufacturer's batches are used for the generation and validation of the substance class *HCK*. The expected variation is also represented in the generation and validation if there are less than three batches.

Also, for the substance class **PhytoComm** (TCM-Granulated herbal extracts of the manufacturer *PhytoComm*) spectra for all useable batches are recorded by *HiperScan GmbH* and entered in the database. The supplier organises the respective tests themself and keeps the test certificates.

A new evaluation option was created for the class *PhytoComm* with the update 2016-01. As the risks are considerably fewer than those from chemical agents, the pharmacist can specify a reasonable criterion for the *specificity* in accordance with internal risk estimation. The database for this is created without taking safety distances into account and no criterion is determined in advance for the *specificity*. Instead, the *specificity* for testing the identity with this concrete substance is calculated in the validation for each substance and provided with the measurement result. The pharmacist then judges himself whether this safety is reasonable with regard to the risk of the substance.

Additionally, a statistical forecast is provided for the *specificity* which is determined according to the *Rule of Three* [10, 11]. For this forecast, it is assumed that there would have been three wrong results more and is provided with a lower limit for *specificity*. This value has a special meaning if a *specificity* of 100 % is achieved for a substance during validation. In this case, the lower limit allows conclusions regarding the scale of existing safety for which with an endless number of validation spectra a value of less than 100 % is to be assumed.

If, for example 14 000 spectra not belonging to the substance are presented and no *false positive* classification is made, a hypothetical number of three *false positive* results is assumed (*Rule of Three* [10, 11]) and the *specificity* is defined with 100.0000 % (> 99.9786 %). Here, it applies that the higher the number of validation spectra which form the statistical basis, the better the *specificity* calculated via the lower *specificity* limit will be approximated.

The positive result of the identity test using *Apo-Ident* establishes that the sample spectrum is in accordance with a batch of the specified granulate from the supplier *PhytoComm*, whereby all useable batches from the supplier are known.

The *PhytoComm* class can only confirm the identity of batches that have been used for the generation of the database. As a consequence there cannot be any validation spectra of other batches. Therefore, the criterion reads that two samples (from different sales containers) from each batch must exist, one for the structure of the database (*Type A*) and one for the validation (*Type B*).

Significance of testing with *Apo-Ident*

The analysis result is determined using sophisticated statistical methods according to state-of-the-art science and technology. Chemical and pharmaceutical knowledge is applied for the selection of the samples from which the calibration spectra and validation spectra are recorded. Otherwise it does not influence the further steps of model creation.

Verbally, the statement of the analysis result can be expressed as follows. Here *"the spectra match"* means that the criteria *Mahalanobis distance*, *outlier analysis* and *correlation* are met as shown in *Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident* [4]. "The spectra do not match", on the other hand, means that at least the criterion *Mahalanobis distance* is not met.

The positive analysis result *"was identified as ..."* is very meaningful because both the quantity of substances to be taken into account and the number of underlying samples is very comprehensive.

1. The spectrum of the sample measured matches spectra of the defined substance.
2. The spectrum of the sample measured does not match any spectrum of any other substance in this substance class. Therefore, all other substances can be clearly ruled-out.
3. As the spectra from other substance classes were used for validation, it is proven that no spectrum of one of these other substances matches the defined substance. (All substance classes with which a spectrum comparison is possible and makes sense are used for validation. This is documented for each substance class in the section *Summary*.)
4. If the defined substance belongs to a *substance group* which in itself is not clearly separable with *Apo-Ident*, matching with the spectra of one or several substances in this group is confirmed. Which of these substances it actually is cannot be determined clearly. All other substances are excluded analogous to 2 and 3.

On the other hand, a negative analysis result *"was not identified as ..."* means:

1. The substance offered could not be recognised based on the spectrum of this sample.
2. The identity of this sample is not confirmed.
3. Testing must be repeated in accordance with the specifications of the Pharmacopoeia.

Conclusion

NIR spectroscopy is a testing method incorporated in the Pharmacopoeia. In case of successful database validation, it is a possible method for identity testing [5]. *Apo-Ident* meets the criteria of the *European Pharmacopeia* as a near-infrared spectrometer and proves the validity of the reference database with the existing validation documentation. This means that *Apo-Ident* can be used as an alternative testing method for testing raw materials at pharmacies.

Explanation of terminology

The following section serves to explain or define specialist terminology which is required in order to understand this document. If necessary, definitions for the analysis system *Apo-Ident* are defined more precisely.

The term database is used in this document exactly as in the *Ph. Eur. Section 2.2.40* [3] synonymous with chemometric model. In order to differentiate the databases which are relatively independent of each other, *HiperScan GmbH* frequently also uses the term substance class (primarily in the plural). On the other hand, the spectra used to structure the database are termed spectrum collection and not database.

Substance classes are units of the organisational structure of the *IdentModule*. The substance classes are substance databases which are also broadly independently subscribable. On the one hand, the liquid and semi-solid substances are separated from the solid powders because they are measured against different references and therefore the spectra cannot be compared. On the other hand, for example the Pharmacopoeia substances are kept separated from the manufacturer-specific database *PhytoComm* for TCM (traditional Chinese medicine) raw materials.

The individual substance classes need only be limited against each other in-part. Often, no risk of confusion exists because they can only be procured from different sources. On the other hand, in several cases we handle substances which need not be distinguished. For example, em Huang Qi granulate from the company *PhytoComm* neither needs to be delimited from *Huang Qi* granulate from the company *HerbaSinica* nor is matching required. Respectively one single chemometric model is behind a substance class. (Even if several reciprocally secured chemometric models would be permissible.) The terms *substance class*, *chemometric model* and *databases* are mostly used here as synonyms.

A substance group respectively summarises all the substances within a substance class which cannot safely be distinguished from one another based on their NIR spectra. However, all the other substances in the database can be excluded.

The formation of subgroups is mentioned in the *Ph. Eur. Section 2.2.40* [3]. In this manner, technical restrictions in case of extensive databases can be avoided and it is possible to prepare individual subgroups with different spectrum pre-treatment. Validation of the subgroups against each other is required. *HiperScan GmbH* has solved these technical restrictions and doesn't use any subgroups within a substance class any longer.

Principal component analysis (PCA) [8, 9] is a multivariate statistics process or multivariate data analysis. It serves to structure, simplify and illustrate comprehensive data records by describing a large number of statistical variables by describing a lower number of linear combinations (the *principal components*) which are as significant as possible. In the *Apo-Ident IdentModule*, *PCA* is used to evaluate the recorded spectrum data (corresponding with *Ph. Eur. 2.2.40* [3]).

The term validation is defined in both relevant contexts here with different (even if related) meanings.

Within the sense of the expert discipline of *chemometrics*, validation is a process step when creating a chemometric model: after a transformation matrix, limits and various parameters have been calculated or determined from a set of reference spectra during the course of the calibration step [8, 9], the validation step determines the performance capability of the model (selectivity, precision, ...) based on the validation spectra. Normally, random sampling is planned here. In order for the validation to gain strength of proof, the validation spectrum set must be selected with an appropriately wide scope (*representative* and *complete*). The terms *validation run* and *validation step* always actually mean the process step in this sense.

In the regulatory sense (of pharmaceutical production), validation is the documented proof that a process or system meets the previously specified requirements reproducibly when applied practi-

cally. In this sense, the *Apo-Ident* databases only become validated databases with the validation documentation, which this document is part of.

The *European Pharmacopeia* uses the term validation in *Section 2.2.40* within the sense of the specialist discipline of *chemometrics* [3].

The robustness of a process is the property of only being influenced by environmental fluctuations (e.g. temperature or humidity) a little. A method is robust if the environmental conditions do not or hardly falsify the final result.

The specificity of a classification (of a chemometric model) is the true negative rate.

The recognition rate (also sensitivity) is the true positive rate. It defines in how many percent of cases a correctly set up substance is actually confirmed.

The true negative rate describes the share of spectra correctly classified as non-identity during validation. This is equivalent to correct classification. It means that a substance A within identity checking as substance B is judged as *"not identified"*. The *true negative rate* is equivalent to the conditional frequency

$$h(\text{rejected}|\text{genuinely no identity}) = \frac{r_n}{r_n + f_p}$$

with r_n as the total number of *true negative* classifications and f_p as the total number of *false positive* classifications. For successful validation of an *IdentModule*, all spectra presented belonging to this category must be classified as *not in accordance*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of this number. The weight of each spectrum of a substance/substance group i therefore results as

$$w_i = \frac{1}{n_i}$$

with n_i number of spectra in this substance/substance group. This weighting ensures that the overall result cannot be enhanced by adding especially large numbers of spectra from easily separable substances.

The true positive rate describes the share of spectra correctly classified as identity during validation. This is equivalent to correct classification. It means that a substance A within identity checking as substance A is judged as *"identified"*. The *true positive rate* is equivalent to the conditional frequency

$$h(\text{identified}|\text{genuine identity}) = \frac{r_p}{r_p + f_n}$$

with r_p as the total number of *true positive* classifications and f_n as the total number of *false negative* classifications. The *true positive* rate is a measure for the recognition rate of the validated *Apo-Ident IdentModule*.
In order to ensure that each substance is received with the same weight, the spectra are weighted as described for the true negative rate.

The true negative result describes a spectrum correctly classified as non-identity during validation. It is equivalent to correct classification. It means that a substance A within identity checking as substance B is judged as *"not identified"*.

The false positive result describes a spectrum falsely classified as non-identity during validation. This is the most critical type of possible false classification. It means that a substance A within identity checking as substance B is judged as *"identified"*. For successful validation of an *IdentModule*, a number of false positive events of zero are demanded for all spectra entering the validation. The exception to this restriction is the class of TCM granulates from the company *PhytoComm* as described under *Particularities of individual substance classes*.

The true positive result describes a spectrum correctly classified as identity during validation. It is equivalent to correct classification. It means that a substance A within identity checking as substance A is judged as *"identified"*.

The false negative result describes a spectrum falsely classified as non-identity during validation. It is equivalent to false classification. It means that a substance A within identity checking as substance A is judged as *"not identified"*.

The 'Rule of Three' says that with a probability of 95 % the next random sample of the same size no more than three false results are to be expected if no false result existed in the existing random sample [10, 11].

The *specificity* and *recognition rate* are determined both globally and from the validation runs for all substances. The information is supplemented with the hypothetical value if there had been three false results more. The percent information is provided in parentheses with the "greater than" symbol '>', e.g. *specificity* 100.000 % (>99.983 %) if 17 567 false spectra have been presented without one single *false positive* result. The larger the statistical basis, the lower the influence of the hypothetical false results.

The Mahalanobis distance is a distance measure between two points in n-dimensional vector space. Here, the respective direction component of the distance to *standard deviation* [12] of an n-dimensional distribution is standardised. In case of the *principal component analysis* [8, 9] this standardisation relates to the distribution of the respective calibration data set for a classification (substance/substance group) in the *principal component space* [8]. The *Mahalanobis distance* of a point (mapping of a spectrum) $\vec{y}$ in the n-dimensional principal component space to the expected value of an n-dimensional distribution $\mathbf{X}$ then results as

$$d(\mathbf{X}, \vec{y}) = \sqrt{(\bar{\mathbf{X}} - \vec{y})^T \mathbf{S}^{-1} (\bar{\mathbf{X}} - \vec{y})} \qquad \text{with} \qquad \mathbf{X} \in \mathbb{R}^{m \times n}, \vec{y} \in \mathbb{R}^m$$

[13]. Here, m is equivalent to the number of principal components used (dimension of the principal component space) and n the number of measurements existing in the calibration data set (spectra). $\bar{\mathbf{X}}$ is the expected value of the resulting distribution for the calibration data set (the average value of n measurements received). $\mathbf{S}^{-1}$ is the inverse covariance matrix [12] for distribution $\mathbf{X}$.

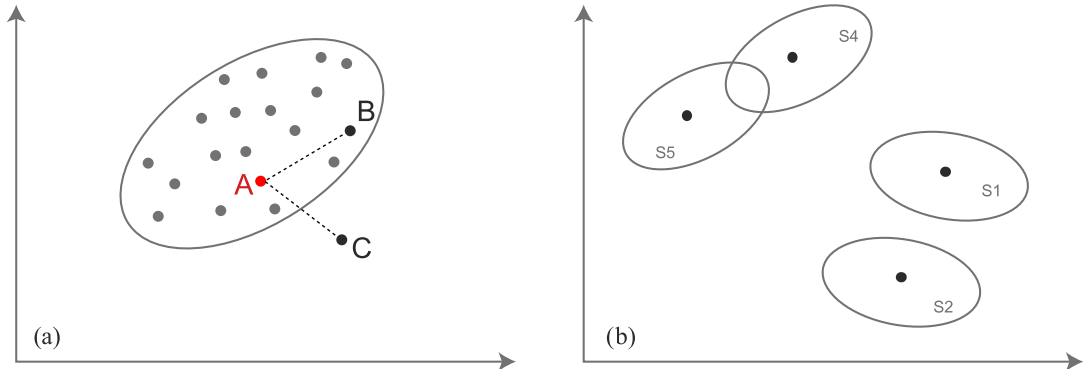

Figure 1: (a) The *Mahalanobis distance* from A to B is less than from A to C. However, the Euclidean distances are identical. (b) The *Mahalanobis distance* between the two measurement series *S4* and *S5* is smaller than between *S1* and *S2*. However, the Euclidean distances are identical.

The *Mahalanobis distance* offers advantages compared to the *Euclidean distance*: For the calculation of the distance it takes the statistical properties of a data point distribution (measurement series), i.e. average value, variance and covariance of the data points [14] into account. The *Mahalanobis distance* is applied while creating the reference database for evaluating the spectra from different samples of a substance.

A chemometric model is a classifier based on statistical methods [8, 9]. Through the respective algorithm used (e.g. *Principal Component Analysis*, *Cluster Analysis*), a maximum of chemical information is extracted from measurement data. Here, systematical or physical disturbances are eliminated using appropriate data pre-processing [15, 16].

At several places in this document, in order to simplify understanding, the term database is used instead of *chemometric model* – in the same manner as in the *Ph. Eur. Section 2.2.40* [3].

A sample (with its own sample ID) refers to substance in a sales container. Repeated sampling from the same sales container is listed under the same sample ID. (The suffix "SI" is not part of the sample ID.) Several samples may originate from the same batch. Samples are called "independent", if they originate from a batch, of which no spectra have been used for the generation of the database. (Up to *IdentModule 2018-01*, samples from different sales packages were considered to be independent.) The information above the list of validation spectra now includes also the number of batches that deliver independent samples for the validation (for both *Type B* and *Type C*).

In case a supplier takes a sample for testing from its incoming goods and splits it to multiple laboratory containers, the substance in all laboratory containers will still be ascribed to the same sample. *HiperScan GmbH* only uses one of the subsamples.

Reference samples are used to structure the database. The *reference spectra* originate from these samples. In chemometric technical jargon you would normally say: For *calibration*, a *chemometric model* is generated from the *calibration spectra* recorded from the *calibration samples*, whose quality is subsequently assessed in validation.

Reference samples are procured via typical pharmacy sources. Their identity is tested. The *reference spectra* are recorded by *HiperScan GmbH*. The documentation also includes the manufacturer's name and batch number.

Reference samples are clearly identified by a sample ID. Samples without sample ID may not be used as *reference samples*.

Summary

A total of 37 934 spectra from 824 different batches for a total of 235 substances were used to validate the substance class *TCM - Granulated herbal extracts (PhytoComm)*.

Validation samples

The validation samples can be categorised as follows:

Typ A Calibration spectra. These are the spectra used to generate the chemometric model. They were recorded by *HiperScan GmbH*. Detailed information regarding the batches or samples can be found in the following validation reports under *calibration samples* and under *Type A*. Further information is listed in *Appendix A*.

Substance class	Substances	Batches	Spectra
TCM - Granulated herbal extracts (PhytoComm)	176	309	24 600

From category *A* a total of 24 600 spectra from 309 batches for a total of 176 substances were taken into account for validation.

Typ B Spectra from independent samples which are not included in database generation. These spectra were recorded by *HiperScan GmbH*. Detailed information regarding the batches or samples can be found in the following validation reports in the section *Type B* or in *Appendix B*.

Substance class	Substances	Batches	Spectra
TCM - Granulated herbal extracts (PhytoComm)	176	309	12 477

From category *B* a total of 12 477 spectra from 309 batches for a total of 176 substances were taken into account for validation.

Typ C Spectra from independent samples which are not included in database generation. *Apo-Ident* customers carried out the measurements. Detailed information regarding the batches or samples can be found in the following validation reports in the section *Type C* or in *Appendix C*.

Substance class	Substances	Batches	Spectra
TCM - Granulated herbal extracts (PhytoComm)	216	519	857

From category *C* a total of 857 spectra from 519 batches for a total of 216 substances were taken into account for validation.

Validation results

The validation runs checked whether the substances/substance groups in the substance class *TCM - Granulated herbal extracts (PhytoComm)* can be distinguished from other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, the matching of all relevant spectra of various substances with the substances/substance groups in the substance class *TCM - Granulated herbal extracts (PhytoComm)* was checked and the correctness of the results was evaluated. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	127	24 367	233	4 304 873
Type B	290	12 132	345	2 183 185
Type C	109	66	630	150 027

Some substances/substance groups in the substance class *TCM - Granulated herbal extracts (PhytoComm)* can only be distinguished from other substances with limitations. (*False positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Typ A	99.996 95 % (> 99.996 86 %)	99.089 13 % (> 99.074 23 %)
Typ B	99.986 09 % (> 99.985 92 %)	97.167 85 % (> 97.138 22 %)
Typ C	99.917 92 % (> 99.914 59 %)	7.557 97 % (> 6.845 57 %)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Bai) Dou Kou**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002257-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Bai) Dou Kou; Amomi cardamomi fructus

Special notes

When selecting the *(Bai) Dou Kou* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Bai) Dou Kou	1	0	2

Second-stage model

For differentiation of the substance/substance group *(Bai) Dou Kou* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Bai) Dou Kou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Bai) Dou Kou	G016H0508921	62761	40	from supplier
PhytoComm	(Bai) Dou Kou	G016H0508921	62762	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *(Bai) Dou Kou*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *(Bai) Dou Kou*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Bai) Dou Kou	G016H0508921	62761†	20
PhytoComm	(Bai) Dou Kou	G016H0508921	62762†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *(Bai) Dou Kou*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	(Bai) Dou Kou	G016H0508521	1
Phytocomm	(Bai) Dou Kou	G016H0508521	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Bai) Dou Kou* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Bai) Dou Kou* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	2	855

The substance/substance group *(Bai) Dou Kou* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Bai) Dou Kou* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Li	12.76	–
Jiang Huang	13.67	–
Sha Ren	13.92	–
He Huan Pi	18.32	–
E Zhu	20.17	–
Sha Shen (Bei)	22.81	–
Mang Xiao	23.88	–
Chen Pi	26.77	–
Di Gu Pi	27.73	–
Yan Hu Suo	31.06	–
Shan Yao	31.16	–
Cang Er Zi	31.46	–
Bai Hua She She Cao	32.56	–
Gan Jiang	32.66	–
Niu Bang Zi	34.02	–
Huang Lian	35.98	–
Zi Su Zi	36.98	–
Du Huo	37.76	–
He Shou Wu	38.75	–
Lai Fu Zi	39.05	–
Wang Bu Liu Xing	40.25	–
Zhi Ke	40.33	–
Wu Yao	40.38	–
Dang Gui	40.41	–
Yi Mu Cao	40.54	–
(Shi) Chang Pu	40.56	–
Xi Xian Cao	40.86	–
Chuan Niu Xi	40.91	–
Sang Ji Shend	41.73	–
Sang Zhi	41.95	–
Xin Yi	42.22	–
Hong Jing Tian	42.31	–
Mi Huan Jun	43.57	–
Hu Zhang	44.19	–
Yin Chen Hao	44.26	–
(Huai) Niu Xi	44.42	–
Chuan Lian Zi	44.83	–
Tu Si Zi	45.07	–
Du Zhong	45.37	–
Mao Dong Qing	45.95	–
Gu Sui Bu	46.17	–
Ze Lan	46.63	–
Bai Xian Pi	46.84	–
Shan Yu Rou	46.90	–
Xiang Fu	47.42	–
Chuan Mu Tong	47.68	–
Xiao Hui Xiang	48.01	–
Tian Hua Fen	48.29	–
Fu Ling	48.98	–
Bu Gu Zhi	49.50	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Bai) Dou Kou* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62761	62761	0.00	12.76
62762	62762	0.00	12.94

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Fen) Bi Xie**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60111-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Fen) Bi Xie; Dioscoreae hypoglaucae rhizoma

Special notes

When selecting the *(Fen) Bi Xie* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Fen) Bi Xie	1	0	1

Second-stage model

For differentiation of the substance/substance group *(Fen) Bi Xie* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Fen) Bi Xie*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Fen) Bi Xie	G290HS286TH1	62815	40	from supplier
PhytoComm	(Fen) Bi Xie	G290HS286TH1	62816	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *(Fen) Bi Xie*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *(Fen) Bi Xie*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Fen) Bi Xie	G290HS286TH1	62815†	20
PhytoComm	(Fen) Bi Xie	G290HS286TH1	62816†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *(Fen) Bi Xie*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	(Fen) Bi Xie	G0290H1251121	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Fen) Bi Xie* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Fen) Bi Xie* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	79	1	24 520
Type B	0	37	3	12 437
Type C	0	0	1	856

The substance/substance group *(Fen) Bi Xie* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9406 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Fen) Bi Xie* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
She Gan	4.32	–
Ren Dong Teng	5.87	–
Lu Gen	6.43	–
Bai Xian Pi	8.24	–
Ye Jiao Teng	8.32	–
Huo Ma Ren	8.39	–
Bai Shao Yao	8.86	–
He Huan Pi	8.89	–
Gu Sui Bu	9.85	–
Rou Gui	9.96	–
Zhu Ling	10.05	–
Gou Teng	10.05	–
Gui Zhi	10.31	–
Ling Zhi	10.60	–
Tu Fu Ling	11.05	–
Fo Shou	11.06	–
Dan Dou Chi	11.16	–
Ji Xue Teng	11.25	–
Lian Zi	11.45	–
Ce Bai Ye	11.77	–
Mu Zei	12.37	–
Ma Huang Gen	12.84	–
Fu Ling	12.95	–
Shen Qu	13.39	–
Guang Huo Xiang	13.49	–
Suan Zao Ren	13.51	–
Tao Ren	14.57	–
Sheng Jiang	14.73	–
Dan Shen	14.81	–
Yu Jin	15.59	–
Yin Yang Huo	15.59	–
Ji Li	15.89	–
Yi Yi Ren	16.07	–
Chuan Xiong	16.26	–
Ci Wu Jia	17.11	–
Ban Xia (Jiang)	17.35	–
Mao Dong Qing	17.51	–
Tai Zi Shen	17.63	–
Gua Lou	17.86	–
Zhe Bei Mu	17.94	–
Fu Zi	18.15	–
Fu Xiao Mai	18.54	–
Ma Huang	18.75	–
Chai Hu	18.78	–
Yan Hu Suo	18.94	–
Shan Yao	19.04	–
Pi Pa Ye	19.29	–
Lian Qiao	20.01	–
Fu Pen Zi	20.56	–
Tian Hua Fen	20.70	–
Gan Cao	20.87	–
Jin Yin Hua	21.45	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jing Jie	21.64	–
Bai Zi Ren	22.53	–
Hong Jing Tian	22.54	–
Chen Pi	22.55	–
Ban Lan Gen	22.75	–
Hou Po	22.78	–
Di Gu Pi	22.92	–
Qiang Huo	23.47	–
Lai Fu Zi	23.65	–
Zhu Ru	23.87	–
Jiao Gu Lan	24.59	–
Fu Shen	25.01	–
Cang Zhu	25.06	–
Ze Xie	25.20	–
Yuan Zhi	25.32	–
Yu Zhu	25.77	–
Zi Hua Di Ding	26.07	–
Che Qian Zi	26.18	–
Sha Ren	27.77	–
Zhi Ke	28.96	–
Dang Gui Wei	29.09	–
Ren Shen	31.42	–
Mang Xiao	31.69	–
Dang Gui	32.10	–
Zhi Gan Cao	32.62	–
Bo He	33.00	–
Jie Geng	33.05	–
Huang Lian	33.08	–
Long Yan Rou	33.47	–
Chuang Mu Xiang	33.79	–
Sang Zhi	34.06	–
Gou Qi Zi	34.20	–
Wu Wei Zi	34.62	–
Huang Qin	34.82	–
San Qi	35.21	–
Du Zhong	35.50	–
Huang Bai	37.24	–
Qing Pi	37.96	–
Zhi Mu	39.49	–
Chi Shao (Yao)	40.35	–
Shan Yu Rou	42.43	–
Mu Gua	43.19	–
Xie Bai	43.65	–
Bai Zhu	44.90	–
Hong Hua	45.36	–
Ku Shen	47.94	–
Bai Jiang Cao	49.11	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Fen) Bi Xie* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62815	62815	0.00	4.32
62816	62816	0.00	4.95

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Huai) Niu Xi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002255-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Huai) Niu Xi; Achyranthis bidentatae radix

Special notes

When selecting the *(Huai) Niu Xi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Huai) Niu Xi	1	0	4

Second-stage model

For differentiation of the substance/substance group *(Huai) Niu Xi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Huai) Niu Xi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Huai) Niu Xi	G003H1905921	62759	40	from supplier
PhytoComm	(Huai) Niu Xi	G003H1905921	62760	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *(Huai) Niu Xi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *(Huai) Niu Xi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Huai) Niu Xi	G003H1905921	62759†	20
PhytoComm	(Huai) Niu Xi	G003H1905921	62760†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 9 spectra from 6 *Apo-Ident* customers from 5 batches from the substance/substance group *(Huai) Niu Xi*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	(Huai) Niu Xi	G003H190522	1
Phytocomm	(Huai) Niu Xi	g003h1905222	1
Phytocomm	(Huai) Niu Xi	G003H1905222	1
PhytoComm	(Huai) Niu Xi	G003H1905222	1
Phytocomm	(Huai) Niu Xi	G003H1905321	3
Phytocomm	(Huai) Niu Xi	G003H1905422	1
PhytoComm	(Huai) Niu Xi	G003H1905422	1

- 848 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Huai) Niu Xi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Huai) Niu Xi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	6	0	9	842

The substance/substance group *(Huai) Niu Xi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.3488 % (> 98.7614 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Huai) Niu Xi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yuan Zhi	11.14	–
Gua Lou	13.63	–
Huang Qi	15.56	–
Jie Geng	15.95	–
Sang Zhi	17.76	–
Bai Zhu	17.79	–
Lian Qiao	18.51	–
Di Gu Pi	20.48	–
Zhi Gan Cao	23.25	–
Chuan Niu Xi	23.88	–
Ku Shen	24.99	–
Tian Hua Fen	25.05	–
Chuan Lian Zi	25.11	–
(Shi) Chang Pu	25.23	–
Chi Shao (Yao)	25.36	–
Zi Su Zi	26.11	–
Niu Bang Zi	26.20	–
Xiang Fu	27.70	–
Du Huo	28.00	–
Mang Xiao	29.21	–
Mu Gua	29.79	–
Shan Yao	30.89	–
Ji Li	30.92	–
Ju Hua	31.05	–
Yan Hu Suo	31.20	–
Lai Fu Zi	31.93	–
Dang Gui	32.80	–
Qin Jiao	34.18	–
Bing Lang	34.27	–
Zhi Ke	35.92	–
Chuan Mu Tong	36.10	–
Xiao Hui Xiang	36.15	–
Ba Ji Tian	36.20	–
E Zhu	36.33	–
Bai Zhi	36.38	–
Chen Pi	36.52	–
Long Dan (Cao)	37.14	–
Xu Duan	37.34	–
Gan Cao	37.34	–
Dong Gua Zi	37.55	–
Bai He	38.26	–
Jin Yin Hua	38.54	–
Cang Er Zi	38.65	–
Wu Yao	38.91	–
Fang Feng	39.03	–
Shan Yu Rou	39.22	–
Sha Shen (Bei)	40.70	–
Gan Jiang	41.01	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pu Gong Ying	41.36	–
Jiang Huang	41.97	–
Lian Zi	42.33	–
Wu Wei Zi	42.53	–
Yi Mu Cao	43.82	–
Fu Zi	44.55	–
Bai Xian Pi	44.73	–
Mi Huan Jun	45.45	–
(Bai) Dou Kou	45.63	–
Da Zao	46.67	–
Sha Ren	46.71	–
Qiang Huo	46.99	–
Chai Hu	48.69	–
Bai Shao Yao	48.78	–
Chuan Xiong	48.80	–
Bai Hua She She Cao	49.57	–
Zhe Bei Mu	49.57	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Huai) Niu Xi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62759	62759	0.00	11.14
62760	62760	0.00	12.81

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Ku) Xing Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50290-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Ku) Xing Ren; Armeniacae amarum semen

Special notes

When selecting the *(Ku) Xing Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Ku) Xing Ren	3	0	5

Second-stage model

For differentiation of the substance/substance group *(Ku) Xing Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Ku) Xing Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Ku) Xing Ren	G029H0708823	62509	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708823	62510	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708921	62767	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708921	62768	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708021	62969	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708021	62970	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *(Ku) Xing Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *(Ku) Xing Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Ku) Xing Ren	G029H0708823	62509†	20
PhytoComm	(Ku) Xing Ren	G029H0708823	62510†	20
PhytoComm	(Ku) Xing Ren	G029H0708921	62767†	20
PhytoComm	(Ku) Xing Ren	G029H0708921	62768†	20
PhytoComm	(Ku) Xing Ren	G029H0708021	62969†	20
PhytoComm	(Ku) Xing Ren	G029H0708021	62970†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 10 spectra from 7 *Apo-Ident* customers from 6 batches from the substance/substance group *(Ku) Xing Ren*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	(Ku) Xing Ren	g029h0708122	1
Phytocomm	(Ku) Xing Ren	G029H0708125	2
PhytoComm	(Ku) Xing Ren	G029H0708125	1
phytocomm	(Ku) Xing Ren	g029h0708322	2
PhytoComm	(Ku) Xing Ren	G029H0708322	2
PhytoComm	(Ku) Xing Ren	G029H0708622	1
PhytoComm	(Ku) Xing Ren	G029H708125	1

- 847 spectra from 13 *Apo-Ident* customers from a total of 513 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Ku) Xing Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Ku) Xing Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	0	10	847

The substance/substance group *(Ku) Xing Ren* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8250 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Ku) Xing Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	50.34	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Ku) Xing Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62509	62509	0.00	54.19
62510	62510	0.00	54.26
62767	62767	0.00	51.24
62768	62768	0.00	51.24
62969	62969	0.00	50.34
62970	62970	0.00	50.40

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Sheng) Di Huang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60005-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Sheng) Di Huang; Rehmanniae radix

Special notes

When selecting the *(Sheng) Di Huang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Sheng) Di Huang	3	0	5

Second-stage model

For differentiation of the substance/substance group *(Sheng) Di Huang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Sheng) Di Huang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Sheng) Di Huang	G211H0532822	62311	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532822	62312	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532921	62749	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532921	62750	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532021	62853	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532021	62854	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *(Sheng) Di Huang*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *(Sheng) Di Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Sheng) Di Huang	G211H0532822	62311†	20
PhytoComm	(Sheng) Di Huang	G211H0532822	62312†	20
PhytoComm	(Sheng) Di Huang	G211H0532921	62749†	20
PhytoComm	(Sheng) Di Huang	G211H0532921	62750†	20
PhytoComm	(Sheng) Di Huang	G211H0532021	62853†	20
PhytoComm	(Sheng) Di Huang	G211H0532021	62854†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 5 *Apo-Ident* customers from 5 batches from the substance/substance group *(Sheng) Di Huang*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
	(Sheng) Di Huang	g211h0532323	1
Phytocomm	(Sheng) Di Huang	g211h0532121	1
Phytocomm	(Sheng) Di Huang	g211h053221	1
Phytocomm	(Sheng) Di Huang	G211H0532521	4
PhytoComm	(Sheng) Di Huang	G211H0532621	1

- 849 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Sheng) Di Huang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Sheng) Di Huang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	2	4	4	847

The substance/substance group *(Sheng) Di Huang* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	99.9380 % (> 99.3506 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Sheng) Di Huang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yu Xing Cao	10.54	–
Xian Mao	14.57	–
Shu Di (Huang)	16.67	–
Mang Xiao	18.16	–
Wu Jia Pi	19.48	–
Ding Xiang	32.01	–
Jing Jie	32.34	–
Guang Huo Xiang	33.79	–
Sang Ye	35.58	–
Dan Zhu Ye	37.42	–
Sang Ji Shend	41.91	–
Hua Shi	45.33	–
Ge Gen	47.61	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Sheng) Di Huang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62311	62311	0.00	10.54
62312	62312	0.00	10.65
62749	62749	0.00	11.81
62750	62750	0.00	12.17
62853	62853	0.00	14.64
62854	62854	0.00	14.82

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**(Shi) Chang Pu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60426-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

(Shi) Chang Pu; Acori rhizoma

Special notes

When selecting the *(Shi) Chang Pu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
(Shi) Chang Pu	2	0	5

Second-stage model

For differentiation of the substance/substance group *(Shi) Chang Pu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *(Shi) Chang Pu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	(Shi) Chang Pu	G006H0544821	62399	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544821	62400	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544922	62945	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544922	62946	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *(Shi) Chang Pu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *(Shi) Chang Pu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	(Shi) Chang Pu	G006H0544821	62399†	20
PhytoComm	(Shi) Chang Pu	G006H0544821	62400†	20
PhytoComm	(Shi) Chang Pu	G006H0544922	62945†	20
PhytoComm	(Shi) Chang Pu	G006H0544922	62946†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 3 *Apo-Ident* customers from 5 batches from the substance/substance group *(Shi) Chang Pu*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Herbasinica	(Shi) Chang Pu	g006h0544321	1
Phytocomm	(Shi) Chang Pu	g006h0544023	1
Phytocomm	(Shi) Chang Pu	G006H0544122	1
Phytocomm	(Shi) Chang Pu	G006H0544322	1
Phytocomm	(Shi) Chang Pu	G006HS270ON1	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *(Shi) Chang Pu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *(Shi) Chang Pu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	1	80	0	12 396
Type C	2	1	4	850

The substance/substance group *(Shi) Chang Pu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9929 % (> 99.9630 %)	100.0000 % (> 92.5000 %)
Type C	99.7674 % (> 99.1806 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *(Shi) Chang Pu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chen Pi	6.37	–
Bai Zhi	6.63	–
Wu Yao	8.30	–
Shan Yu Rou	8.38	–
Chuan Niu Xi	12.88	–
Xiang Fu	13.71	–
Dang Gui	15.97	–
Ju Hua	17.67	–
Sha Ren	18.19	–
Tian Hua Fen	19.19	–
Mu Dan Pi	19.30	–
Gua Lou	19.85	–
Du Huo	21.68	–
Zhi Gan Cao	21.94	–
Gu Sui Bu	21.98	–
Ji Li	22.28	–
He Huan Pi	23.07	–
Xiao Hui Xiang	23.11	–
Jiang Huang	23.20	–
Huang Qi	23.83	–
Chuan Lian Zi	23.85	–
Yan Hu Suo	23.93	–
Di Gu Pi	24.17	–
Wu Zhu Yu	24.21	–
(Huai) Niu Xi	24.61	–
Zi Su Zi	25.18	–
Fang Feng	25.49	–
Shan Yao	25.64	–
Yi Mu Cao	25.77	–
Hong Jing Tian	25.83	–
Mi Huan Jun	26.09	–
Xu Duan	26.57	–
Mang Xiao	27.71	–
Ku Shen	28.46	–
Bai Hua She She Cao	28.75	–
Sang Zhi	28.81	–
Sang Ji Shend	29.02	–
(Bai) Dou Kou	29.10	–
Jie Geng	30.02	–
Ba Ji Tian	30.06	–
Bing Lang	30.55	–
E Zhu	31.76	–
Niu Bang Zi	31.84	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	32.30	–
Cang Er Zi	32.66	–
Chi Shao (Yao)	32.83	–
Lian Qiao	32.91	–
Pu Gong Ying	33.83	–
Gan Jiang	34.50	–
Sang Bai Pi	35.15	–
Yin Chen Hao	35.36	–
Xin Yi	35.60	–
Bai Zhu	36.05	–
He Shou Wu	36.14	–
Wu Wei Zi	37.19	–
Huang Lian	37.39	–
Qin Jiao	38.71	–
Sha Shen (Bei)	38.99	–
Sang Ye	39.47	–
Mao Dong Qing	39.62	–
Bo He	39.69	–
Ge Gen	40.39	–
Ze Lan	40.54	–
Lai Fu Zi	40.59	–
Chuan Mu Tong	41.77	–
Nü Zhen Zi	42.25	–
Du Zhong	42.25	–
Tu Fu Ling	43.04	–
Jin Yin Hua	43.10	–
Dan Zhu Ye	43.14	–
Yuan Zhi	43.75	–
Suan Zao Ren	43.83	–
Gan Cao	44.02	–
Tu Si Zi	44.57	–
Bai He	45.06	–
Hu Zhang	45.47	–
Xi Xian Cao	46.67	–
Da Zao	46.73	–
Qiang Huo	46.89	–
Lian Zi	46.95	–
Zi Hua Di Ding	47.02	–
Dong Gua Zi	47.34	–
Bai Jiang Cao	48.25	–
Bai Xian Pi	49.15	–
Cang Zhu	49.82	–
Bu Gu Zhi	49.90	–
Jing Jie	50.05	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *(Shi) Chang Pu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62399	62399	0.00	6.63
62400	62400	0.00	7.38
62945	62945	0.00	6.37
62946	62946	0.00	6.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ba Ji Tian**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60138-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ba Ji Tian; Morindae officinalis radix

Special notes

When selecting the *Ba Ji Tian* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ba Ji Tian	1	0	0

Second-stage model

For differentiation of the substance/substance group *Ba Ji Tian* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ba Ji Tian*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ba Ji Tian	G165H0445921	62823	40	from supplier
PhytoComm	Ba Ji Tian	G165H0445921	62824	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ba Ji Tian*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ba Ji Tian*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ba Ji Tian	G165H0445921	62823†	20
PhytoComm	Ba Ji Tian	G165H0445921	62824†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ba Ji Tian*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ba Ji Tian* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ba Ji Tian* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	0	857

The substance/substance group *Ba Ji Tian* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ba Ji Tian* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
(Shi) Chang Pu	13.35	–
Chen Pi	13.37	–
Shan Yu Rou	16.42	–
Mu Dan Pi	16.59	–
Chuan Niu Xi	18.00	–
Chuan Lian Zi	18.67	–
He Huan Pi	20.97	–
Sang Zhi	22.52	–
Fang Feng	23.43	–
Ji Li	23.56	–
Mang Xiao	23.74	–
Ju Hua	23.94	–
Dang Gui	24.08	–
Wu Yao	24.40	–
Xiang Fu	24.56	–
Jiang Huang	27.03	–
(Huai) Niu Xi	27.56	–
E Zhu	28.53	–
Long Dan (Cao)	28.96	–
Du Huo	29.99	–
Gu Sui Bu	31.59	–
Da Zao	31.79	–
Lai Fu Zi	32.19	–
Sha Ren	33.20	–
Huang Qi	33.38	–
Yan Hu Suo	34.11	–
Zhi Gan Cao	34.20	–
Bing Lang	34.56	–
Hong Jing Tian	36.93	–
Shan Yao	37.66	–
Sang Ji Shend	38.57	–
Zi Su Zi	38.69	–
Yin Chen Hao	40.08	–
Mai Ya	41.39	–
Xiao Hui Xiang	41.71	–
(Bai) Dou Kou	42.23	–
Dong Gua Zi	42.62	–
Niu Bang Zi	42.97	–
Tian Hua Fen	43.34	–
Di Gu Pi	43.59	–
Bai Zhu	44.80	–
Jie Geng	45.34	–
Yuan Zhi	45.64	–
Sang Bai Pi	46.29	–
Sha Shen (Bei)	46.49	–
Gua Lou	46.68	–
Suan Zao Ren	47.26	–
He Shou Wu	47.31	–
Xu Duan	47.39	–
Gan Jiang	47.39	–
Tu Si Zi	47.54	–
Mi Huan Jun	47.66	–
Xin Yi	48.06	–
Bai Hua She She Cao	48.19	–
Qin Jiao	48.71	–
Zhi Ke	48.74	–
Wu Zhu Yu	49.58	–
Ku Shen	49.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ba Ji Tian* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62823	62823	0.00	13.37
62824	62824	0.00	13.35

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai He**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60093-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai He; Lilii bulbus

Special notes

When selecting the *Bai He* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai He	2	0	2

Second-stage model

For differentiation of the substance/substance group *Bai He* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai He*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai He	G139H0607921	62739	40	from supplier
PhytoComm	Bai He	G139H0607921	62740	40	from supplier
PhytoComm	Bai He	G139H0607922	62887	40	from supplier
PhytoComm	Bai He	G139H0607922	62888	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Bai He*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Bai He*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai He	G139H0607921	62739†	20
PhytoComm	Bai He	G139H0607921	62740†	20
PhytoComm	Bai He	G139H0607922	62887†	20
PhytoComm	Bai He	G139H0607922	62888†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 5 *Apo-Ident* customers from 3 batches from the substance/substance group *Bai He*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai He	g139h0607221	1
PhytoComm	Bai He	G139H0607221	1
Phytocomm	Bai He	G139H0607421	2
PhytoComm	Bai He	G139H0607421	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai He* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai He* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	5	852

The substance/substance group *Bai He* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai He* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sang Zhi	9.78	–
Chuan Lian Zi	13.28	–
Mai Ya	15.07	–
Sha Shen (Bei)	16.53	–
Lai Fu Zi	17.15	–
Long Dan (Cao)	18.37	–
Da Zao	18.38	–
Di Gu Pi	19.15	–
Dang Gui	19.84	–
Tian Hua Fen	20.56	–
Mi Huan Jun	21.61	–
Mang Xiao	23.21	–
Chuan Mu Tong	23.82	–
Gan Jiang	24.34	–
Jie Geng	25.08	–
Bai Zhu	27.06	–
Qin Jiao	27.35	–
E Zhu	27.58	–
Huang Qi	28.43	–
Zhi Gan Cao	34.60	–
Yan Hu Suo	34.61	–
Zi Su Zi	36.20	–
Mu Gua	36.26	–
Shan Yao	38.40	–
Jiang Huang	38.45	–
Cang Er Zi	38.80	–
(Huai) Niu Xi	38.85	–
Ban Zhi Lian	39.45	–
Suan Zao Ren	39.80	–
Ji Li	40.73	–
Yi Yi Ren	41.61	–
(Shi) Chang Pu	41.72	–
Rou Gui	42.00	–
Chi Shao (Yao)	43.49	–
Fu Ling	43.87	–
Lian Zi	44.01	–
Tu Fu Ling	45.15	–
Bai Xian Pi	45.34	–
Bai Zhi	45.72	–
Chuan Niu Xi	45.78	–
Sheng Jiang	46.04	–
Lian Qiao	46.65	–
Gua Lou	47.41	–
Yuan Zhi	47.60	–
Bing Lang	48.20	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chai Hu	48.79	–
Ma Huang Gen	49.23	–
Bai Shao Yao	49.89	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai He* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62739	62739	0.00	9.78
62740	62740	0.00	10.49
62887	62887	0.00	14.43
62888	62888	0.00	13.28

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Hua She She Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60039-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Hua She She Cao; Oldenlandiae diffusae herba

Special notes

When selecting the *Bai Hua She She Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Hua She She Cao	2	0	2

Second-stage model

For differentiation of the substance/substance group *Bai Hua She She Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Hua She She Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Hua She She …	G173HS078RW1	62575	40	from supplier
PhytoComm	Bai Hua She She …	G173HS078RW1	62576	40	from supplier
PhytoComm	Bai Hua She She …	G173HS078SP1	62827	40	from supplier
PhytoComm	Bai Hua She She …	G173HS078SP1	62828	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Bai Hua She She Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Bai Hua She She Cao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Hua She She Cao	G173HS078RW1	62575†	20
PhytoComm	Bai Hua She She Cao	G173HS078RW1	62576†	20
PhytoComm	Bai Hua She She Cao	G173HS078SP1	62827†	20
PhytoComm	Bai Hua She She Cao	G173HS078SP1	62828†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 7 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Bai Hua She She Cao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Hua She She Cao	G173H1153222	3
Phytocomm	Bai Hua She She Cao	G173H1153421	4

- 850 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Hua She She Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Hua She She Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	1	80	0	12 396
Type C	0	0	7	850

The substance/substance group *Bai Hua She She Cao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9952 % (> 99.9654 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8255 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Hua She She Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Xin Yi	6.89	–
Han Lian Cao	9.78	–
Zhi Shi	10.33	–
Xi Xian Cao	11.27	–
Xia Ku Cao	12.11	–
Yu Xing Cao	12.66	–
Huang Lian	14.11	–
Hu Zhang	14.44	–
Jiang Huang	15.62	–
Sang Ye	16.45	–
Tu Fu Ling	19.49	–
Ze Lan	20.45	–
Pu Gong Ying	21.71	–
Shan Yao	21.88	–
Xiang Fu	22.92	–
Sha Ren	24.89	–
Yi Mu Cao	26.80	–
Mang Xiao	27.54	–
He Huan Pi	27.60	–
Nü Zhen Zi	28.29	–
Rou Cong Rong	30.67	–
Yin Chen Hao	31.24	–
Du Zhong	33.77	–
Gu Sui Bu	34.57	–
Sang Ji Shend	35.06	–
Gou Teng	36.99	–
Zhi Ke	37.46	–
Jiao Gu Lan	37.51	–
E Zhu	38.70	–
Yu Jin	39.00	–
Dan Shen	39.01	–
Ji Li	40.11	–
Yan Hu Suo	40.42	–
Cang Er Zi	41.27	–
Fu Zi	41.82	–
(Bai) Dou Kou	42.31	–
Qiang Huo	42.69	–
Bai Jiang Cao	43.70	–
(Shi) Chang Pu	43.98	–
Che Qian Zi	44.41	–
Bo He	44.45	–
Chen Pi	45.66	–
Niu Bang Zi	46.62	–
Ju Hua	46.84	–
Bai Xian Pi	47.46	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gua Lou	47.56	–
Wu Yao	47.58	–
Mao Dong Qing	48.02	–
Hou Po	48.64	–
Wu Wei Zi	48.69	–
Dan Zhu Ye	49.38	–
Wu Mei	49.73	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Hua She She Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62575	62575	0.00	6.89
62576	62576	0.00	7.50
62827	62827	0.00	11.91
62828	62828	0.00	11.27

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Jiang Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50319-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Jiang Cao; Patriniae herba

Special notes

When selecting the *Bai Jiang Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Jiang Cao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Bai Jiang Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Jiang Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Jiang Cao	G181HS381TH1	62833	40	from supplier
PhytoComm	Bai Jiang Cao	G181HS381TH1	62834	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Bai Jiang Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Bai Jiang Cao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Jiang Cao	G181HS381TH1	62833†	20
PhytoComm	Bai Jiang Cao	G181HS381TH1	62834†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Bai Jiang Cao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Jiang Cao	G181H1180321	1
Phytocomm	Bai Jiang Cao	G181H1180422	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Jiang Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Jiang Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	39	1	12 437
Type C	0	0	2	855

The substance/substance group *Bai Jiang Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Jiang Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sha Ren	6.72	–
Du Zhong	9.50	–
Jiao Gu Lan	9.89	–
Bo He	10.41	–
Yu Jin	11.29	–
Mao Dong Qing	12.06	–
Pu Gong Ying	12.43	–
Fu Zi	12.89	–
Dan Shen	13.05	–
Chai Hu	13.14	–
Qiang Huo	13.30	–
Tu Fu Ling	13.41	–
Zhi Ke	13.61	–
Huang Lian	13.87	–
Huang Bai	15.63	–
Qing Hao	15.93	–
Ze Lan	16.56	–
Zi Hua Di Ding	16.74	–
Hou Po	18.23	–
Che Qian Zi	18.37	–
Yin Yang Huo	18.92	–
Bai Xian Pi	19.05	–
Shan Yao	19.40	–
Jing Jie	19.42	–
Ye Jiao Teng	20.22	–
Xi Xian Cao	20.50	–
Qing Pi	20.83	–
Sang Ye	21.95	–
Fu Pen Zi	22.23	–
Hong Jing Tian	22.37	–
Ce Bai Ye	22.40	–
Dan Dou Chi	22.98	–
Jin Yin Hua	23.27	–
Ling Zhi	23.46	–
Pi Pa Ye	24.59	–
Ma Huang	25.19	–
Wu Wei Zi	25.21	–
Yan Hu Suo	25.60	–
Zhi Shi	26.58	–
Di Gu Pi	26.66	–
Ji Xue Teng	27.02	–
Nü Zhen Zi	27.06	–
Chuan Xiong	27.27	–
Lian Zi	27.40	–
Ren Dong Teng	28.52	–
Chuang Mu Xiang	28.62	–
Tian Hua Fen	28.68	–
Bai Shao Yao	29.77	–
Mang Xiao	30.32	–
Yin Chen Hao	30.34	–
Gu Sui Bu	30.79	–
He Huan Pi	31.18	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shen Qu	31.69	–
Suan Zao Ren	31.91	–
Wu Mei	33.09	–
Gan Cao	33.33	–
Ji Li	34.22	–
Ren Shen	34.40	–
Xin Yi	34.43	–
Ban Lan Gen	34.63	–
Jiang Huang	34.85	–
Lian Qiao	34.95	–
Cang Er Zi	35.25	–
Guang Huo Xiang	35.63	–
(Fen) Bi Xie	37.21	–
Zhu Ling	37.69	–
Lu Gen	37.78	–
Hong Hua	37.97	–
Gou Teng	38.80	–
Gou Qi Zi	38.85	–
Gua Lou	39.40	–
Yuan Zhi	39.50	–
E Zhu	39.90	–
Mu Zei	40.53	–
Cang Zhu	40.60	–
Chi Shao (Yao)	41.03	–
Yi Mu Cao	41.31	–
Zhe Bei Mu	41.36	–
Sheng Jiang	41.68	–
Zhi Gan Cao	42.73	–
Bai Hua She She Cao	42.73	–
She Gan	43.05	–
Rou Gui	43.11	–
Ban Zhi Lian	43.34	–
Tao Ren	43.94	–
Zi Su Zi	44.04	–
Niu Bang Zi	44.17	–
Shan Yu Rou	44.25	–
Huang Qin	44.60	–
Wang Bu Liu Xing	44.89	–
(Shi) Chang Pu	45.53	–
Sha Shen (Bei)	45.70	–
Jie Geng	46.24	–
Ma Huang Gen	46.26	–
Yu Xing Cao	47.43	–
Hu Zhang	47.61	–
Huo Ma Ren	48.73	–
Tai Zi Shen	50.06	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Jiang Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62833	62833	0.00	7.35
62834	62834	0.00	6.72

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Shao Yao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60007-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Shao Yao; Paeoniae lactiflorae albus radix

Special notes

When selecting the *Bai Shao Yao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Shao Yao	4	0	5

Second-stage model

For differentiation of the substance/substance group *Bai Shao Yao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Shao Yao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Shao Yao	G179HS065RQ1	62275	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065RT1	62435	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065RT1	62436	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065RW1	62601	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065RW1	62602	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065TK1	62829	40	from supplier
PhytoComm	Bai Shao Yao	G179HS065TK1	62830	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 280 spectra of 7 reference samples from the substance/substance group *Bai Shao Yao*. These samples are listed above in the *calibration samples* section. The reference samples come from 4 different batches.

- 24 320 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 199 spectra of 8 reference samples from the substance/substance group *Bai Shao Yao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Shao Yao	G179HS065RQ1	62275†	20
PhytoComm	Bai Shao Yao	G179HS065RQ1	62276	59
PhytoComm	Bai Shao Yao	G179HS065RT1	62435†	20
PhytoComm	Bai Shao Yao	G179HS065RT1	62436†	20
PhytoComm	Bai Shao Yao	G179HS065RW1	62601†	20

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Shao Yao	G179HS065RW1	62602†	20
PhytoComm	Bai Shao Yao	G179HS065TK1	62829†	20
PhytoComm	Bai Shao Yao	G179HS065TK1	62830†	20

- 12 278 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 11 spectra from 6 *Apo-Ident* customers from 6 batches from the substance/substance group *Bai Shao Yao*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Bai Shao Yao	1214401	1
Phytocomm	Bai Shao Yao	G179H0502121	1
Phytocomm	Bai Shao Yao	g179h0502322	1
Phytocomm	Bai Shao Yao	G179H0502322	1
PhytoComm	Bai Shao Yao	G179H0502322	1
Phytocomm	Bai Shao Yao	G179H0502521	5
Phytocomm	Bai Shao Yao	g17h0502323	1

- 846 spectra from 13 *Apo-Ident* customers from a total of 513 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Shao Yao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Shao Yao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	2	280	0	24 318
Type B	3	188	11	12 275
Type C	1	0	11	845

†These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

The substance/substance group *Bai Shao Yao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9893 % (> 99.9742 %)	100.0000 % (> 97.8571 %)
Type B	99.9833 % (> 99.9534 %)	94.4724 % (> 92.9648 %)
Type C	99.5349 % (> 98.9473 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Shao Yao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ban Xia (Jiang)	3.08	–
Tai Zi Shen	5.74	–
Di Gu Pi	6.23	–
Ye Jiao Teng	7.01	–
He Huan Pi	7.72	–
Zhu Ru	7.86	–
Rou Gui	7.93	–
Gua Lou	7.94	–
Pi Pa Ye	7.99	–
Zhe Bei Mu	8.62	–
Yuan Zhi	8.95	–
Gui Zhi	9.86	–
Ji Li	9.91	–
Shen Qu	10.00	–
Ci Wu Jia	10.02	–
Sheng Jiang	10.11	–
Shan Yao	10.36	–
Fu Xiao Mai	10.66	–
Tian Hua Fen	10.67	–
Fo Shou	11.80	–
Bai Zi Ren	11.97	–
Tao Ren	12.17	–
Lai Fu Zi	12.24	–
Ji Xue Teng	13.40	–
Yi Yi Ren	13.92	–
Gou Teng	14.40	–
Jin Yin Hua	14.43	–
Mu Zei	14.54	–
Lian Zi	15.05	–
Lian Qiao	15.11	–
Zhi Mu	15.29	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Ling	15.29	–
Huo Ma Ren	16.88	–
Ling Zhi	17.19	–
San Qi	18.31	–
Ren Dong Teng	18.52	–
Dang Gui Wei	18.65	–
Ze Xie	19.23	–
Lu Gen	19.62	–
Dang Gui	19.86	–
Ku Shen	20.08	–
Bai Xian Pi	20.16	–
Fu Zi	20.35	–
Ren Shen	20.36	–
Shan Yu Rou	20.41	–
Gu Sui Bu	20.42	–
Jie Geng	20.93	–
Chen Pi	21.45	–
Fu Shen	21.72	–
Ma Huang Gen	22.25	–
Chuan Xiong	22.30	–
Yu Zhu	22.62	–
Cang Zhu	23.25	–
Ban Lan Gen	24.49	–
Jiao Gu Lan	24.50	–
Yan Hu Suo	24.72	–
Zhi Ke	24.92	–
Fu Pen Zi	25.16	–
Tu Fu Ling	25.30	–
Suan Zao Ren	25.52	–
Yu Jin	25.79	–
Huang Qin	26.63	–
Mu Gua	27.08	–
Gou Qi Zi	27.57	–
Wu Wei Zi	27.61	–
Hou Po	28.39	–
Ma Huang	29.28	–
Che Qian Zi	29.31	–
Zhu Ling	29.77	–
(Fen) Bi Xie	30.10	–
Dan Shen	30.25	–
Yin Yang Huo	31.29	–
Ce Bai Ye	31.67	–
Dan Dou Chi	31.98	–
Mang Xiao	32.30	–
Mao Dong Qing	32.35	–
Guang Huo Xiang	34.02	–
Long Yan Rou	35.75	–
Chai Hu	36.04	–
She Gan	36.41	–
Mai Men Dong	36.84	–
Xie Bai	37.31	–
Gan Cao	38.47	–
Hong Jing Tian	38.69	–
Qiang Huo	40.63	–
Chuang Mu Xiang	41.77	–
Sang Zhi	43.26	–
Huang Lian	45.65	–
Qing Pi	46.10	–
Chi Shao (Yao)	46.74	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	49.34	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Shao Yao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62275	62275	0.00	5.74
62435	62435	0.00	8.20
62436	62436	0.00	7.86
62601	62601	0.00	9.85
62602	62602	0.00	9.91
62829	62829	0.00	3.11
62830	62830	0.00	3.08

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Xian Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60208-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Xian Pi; Dictamni cortex

Special notes

When selecting the *Bai Xian Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Xian Pi	2	0	1

Second-stage model

For differentiation of the substance/substance group *Bai Xian Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Xian Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Xian Pi	G090HS335RP1	62285	40	from supplier
PhytoComm	Bai Xian Pi	G090HS335RP1	62286	40	from supplier
PhytoComm	Bai Xian Pi	G090HS335RN1	62449	40	from supplier
PhytoComm	Bai Xian Pi	G090HS335RN1	62450	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Bai Xian Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Bai Xian Pi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Xian Pi	G090HS335RP1	62285†	20
PhytoComm	Bai Xian Pi	G090HS335RP1	62286†	20
PhytoComm	Bai Xian Pi	G090HS335RN1	62449†	20
PhytoComm	Bai Xian Pi	G090HS335RN1	62450†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Bai Xian Pi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Xian Pi	G090H0515121	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Xian Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Xian Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	153	7	24 437
Type B	7	80	0	12 390
Type C	1	0	1	855

The substance/substance group *Bai Xian Pi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9893 % (> 99.9743 %)	95.6250 % (> 93.7500 %)
Type B	99.9548 % (> 99.9249 %)	100.0000 % (> 92.5000 %)
Type C	99.9225 % (> 99.3408 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Xian Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dan Dou Chi	3.49	–
Ye Jiao Teng	5.43	–
Tu Fu Ling	6.25	–
Ling Zhi	7.05	–
Yin Yang Huo	7.82	–
Mao Dong Qing	8.56	–
Ren Dong Teng	9.30	–
Yu Jin	9.55	–
Ce Bai Ye	9.62	–
Hou Po	10.62	–
Jing Jie	10.68	–
Chai Hu	11.15	–
Lian Zi	11.47	–
Fu Zi	12.39	–
Shen Qu	12.64	–
Zhu Ling	12.74	–
Che Qian Zi	13.08	–
Qiang Huo	13.15	–
Guang Huo Xiang	13.89	–
Ji Xue Teng	14.60	–
Zi Hua Di Ding	14.79	–
Rou Gui	14.96	–
(Fen) Bi Xie	15.07	–
Dan Shen	16.26	–
He Huan Pi	17.58	–
Sha Ren	17.86	–
Gou Teng	18.12	–
Lu Gen	18.39	–
Suan Zao Ren	18.55	–
Yan Hu Suo	18.98	–
Ma Huang Gen	19.55	–
Shan Yao	19.75	–
Ji Li	19.98	–
Du Zhong	19.99	–
Gan Cao	20.21	–
Bo He	20.38	–
Fu Pen Zi	20.44	–
Pi Pa Ye	20.89	–
Lian Qiao	20.89	–
Jin Yin Hua	21.03	–
Gu Sui Bu	21.28	–
Chuan Xiong	21.51	–
Huo Ma Ren	21.52	–
Huang Bai	21.58	–
Ma Huang	21.80	–
Sheng Jiang	21.80	–
Zhi Ke	22.55	–
She Gan	22.96	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	23.23	–
Bai Shao Yao	23.40	–
Gui Zhi	23.63	–
Tian Hua Fen	24.85	–
Yi Yi Ren	25.02	–
Qing Pi	25.32	–
Di Gu Pi	25.42	–
Fu Ling	25.49	–
Hong Jing Tian	25.56	–
Chuang Mu Xiang	26.14	–
Mu Zei	26.41	–
Gua Lou	26.46	–
Huang Lian	27.57	–
Tao Ren	27.77	–
Qing Hao	28.37	–
Ban Lan Gen	28.55	–
Wu Wei Zi	28.95	–
Zhe Bei Mu	29.08	–
Tai Zi Shen	29.27	–
Zhi Gan Cao	30.18	–
Fo Shou	30.75	–
Cang Zhu	31.22	–
Bai Jiang Cao	31.28	–
Yuan Zhi	31.35	–
Mang Xiao	31.75	–
Ci Wu Jia	32.52	–
Ban Zhi Lian	33.45	–
Fu Shen	34.05	–
Gou Qi Zi	34.57	–
Ban Xia (Jiang)	35.10	–
Fu Xiao Mai	35.19	–
Ren Shen	35.57	–
Chen Pi	35.96	–
Bai Zi Ren	36.77	–
Yu Zhu	37.47	–
Sang Zhi	38.40	–
Zhu Ru	39.92	–
Chi Shao (Yao)	40.58	–
Dang Gui Wei	41.46	–
Huang Qin	42.65	–
Jie Geng	42.74	–
Dang Gui	43.79	–
Ze Xie	44.17	–
Pu Gong Ying	44.23	–
Hong Hua	45.24	–
San Qi	45.95	–
Ze Lan	46.38	–
Xie Bai	46.44	–
Bai Zhu	47.14	–
Shan Yu Rou	48.13	–
Cang Er Zi	48.40	–
Long Yan Rou	48.88	–
Lai Fu Zi	48.90	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Xian Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62285	62285	0.00	3.49
62286	62286	0.00	3.56
62449	62449	0.00	4.80
62450	62450	0.00	4.56

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60235-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Zhi; Angelicae dahuricae radix

Special notes

When selecting the *Bai Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Zhi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Bai Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Zhi	G020H0503922	62961	40	from supplier
PhytoComm	Bai Zhi	G020H0503922	62962	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Bai Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Bai Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Zhi	G020H0503922	62961†	20
PhytoComm	Bai Zhi	G020H0503922	62962†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Bai Zhi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Zhi	G020H0503421	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	1	0	856

The substance/substance group *Bai Zhi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
(Shi) Chang Pu	8.71	–
Chuan Niu Xi	13.68	–
Xiang Fu	15.64	–
Chen Pi	16.25	–
Dang Gui	16.77	–
Du Huo	20.92	–
Wu Yao	21.82	–
Tian Hua Fen	22.00	–
Gua Lou	22.26	–
Zhi Gan Cao	24.36	–
Ju Hua	25.55	–
Shan Yu Rou	25.70	–
Yan Hu Suo	25.93	–
Huang Qi	26.74	–
(Huai) Niu Xi	27.07	–
Chuan Lian Zi	27.08	–
Shan Yao	27.18	–
Sha Ren	27.29	–
Mang Xiao	27.46	–
Di Gu Pi	27.55	–
Xu Duan	27.62	–
Ku Shen	27.87	–
Zi Su Zi	28.70	–
Mu Dan Pi	28.72	–
Mi Huan Jun	29.26	–
Bai Hua She She Cao	29.86	–
Sang Zhi	30.13	–
Bing Lang	30.96	–
Jiang Huang	31.18	–
Yi Mu Cao	31.24	–
Gu Sui Bu	31.24	–
(Bai) Dou Kou	31.50	–
Lian Qiao	31.77	–
Wu Zhu Yu	32.09	–
Xiao Hui Xiang	32.21	–
Gan Jiang	32.54	–
Fang Feng	34.16	–
Chi Shao (Yao)	34.29	–
Sang Ye	34.45	–
Pu Gong Ying	34.54	–
Cang Er Zi	35.26	–
Zhi Ke	35.41	–
Jie Geng	35.44	–
Niu Bang Zi	35.61	–
Ji Li	35.74	–
Xin Yi	36.24	–
Sang Ji Shend	36.70	–
Ba Ji Tian	37.03	–
Bai Zhu	37.91	–
Wu Wei Zi	38.38	–
Qin Jiao	38.91	–
He Huan Pi	39.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	39.51	–
Bo He	40.02	–
Ze Lan	40.81	–
Chuan Mu Tong	42.21	–
Sha Shen (Bei)	42.59	–
Mao Dong Qing	42.70	–
Huang Lian	42.96	–
Gan Cao	42.98	–
He Shou Wu	42.99	–
Tu Fu Ling	44.08	–
Du Zhong	46.01	–
Ge Gen	46.17	–
Hong Jing Tian	46.37	–
Yuan Zhi	46.64	–
Jin Yin Hua	46.93	–
Qiang Huo	47.17	–
Xi Xian Cao	47.26	–
Zi Hua Di Ding	47.46	–
Lai Fu Zi	48.01	–
Nü Zhen Zi	49.90	–
Hu Zhang	50.04	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Zhi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62961	62961	0.00	9.40
62962	62962	0.00	8.71

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Zhu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60015-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Zhu; Atractylodis macrocephalae rhizoma

Special notes

When selecting the *Bai Zhu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Zhu	2	0	3

Second-stage model

For differentiation of the substance/substance group *Bai Zhu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Zhu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Zhu	G041H0501822	62385	40	from supplier
PhytoComm	Bai Zhu	G041H0501822	62386	40	from supplier
PhytoComm	Bai Zhu	G041H0501922	62975	40	from supplier
PhytoComm	Bai Zhu	G041H0501922	62976	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Bai Zhu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Bai Zhu*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Zhu	G041H0501822	62385†	20
PhytoComm	Bai Zhu	G041H0501822	62386†	20
PhytoComm	Bai Zhu	G041H0501922	62975†	20
PhytoComm	Bai Zhu	G041H0501922	62976†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Bai Zhu*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Zhu	g041h0501221	3
Phytocomm	Bai Zhu	G041H0501221	1
Phytocomm	Bai Zhu	G041H0501422	3
PhytoComm	Bai Zhu	G041H0501621	1

- 849 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Zhu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Zhu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	79	1	12 397
Type C	3	1	7	846

The substance/substance group *Bai Zhu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	98.7500 % (> 95.0000 %)
Type C	99.5814 % (> 98.9940 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Zhu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jie Geng	5.95	–
Mu Gua	10.74	–
Zhi Gan Cao	11.65	–
Huang Qi	12.29	–
Yuan Zhi	14.25	–
(Huai) Niu Xi	15.12	–
Sang Zhi	16.86	–
Chuan Lian Zi	16.92	–
Di Gu Pi	17.29	–
Long Dan (Cao)	20.68	–
Tian Hua Fen	21.53	–
Qin Jiao	22.41	–
Chi Shao (Yao)	22.49	–
Gua Lou	22.65	–
Bai He	23.18	–
Lai Fu Zi	24.06	–
Ji Li	28.92	–
Mang Xiao	29.35	–
Zi Su Zi	30.04	–
Sha Shen (Bei)	30.10	–
Lian Qiao	30.92	–
Chuan Mu Tong	31.01	–
Dang Gui	31.55	–
Gan Jiang	32.05	–
Shan Yao	32.90	–
Bai Xian Pi	32.92	–
Tu Fu Ling	32.93	–
Bing Lang	33.44	–
Chuan Niu Xi	33.69	–
E Zhu	34.00	–
Lian Zi	35.22	–
(Shi) Chang Pu	35.94	–
Ku Shen	36.43	–
Bai Shao Yao	36.46	–
Bai Zhi	36.65	–
Niu Bang Zi	37.20	–
Ban Zhi Lian	37.66	–
Du Huo	38.30	–
Chai Hu	39.03	–
Rou Gui	39.14	–
Cang Er Zi	39.49	–
Fu Ling	40.63	–
Mi Huan Jun	40.90	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	41.06	–
Mai Ya	41.13	–
Zhi Ke	41.56	–
Ju Hua	42.18	–
Suan Zao Ren	42.33	–
Ye Jiao Teng	42.89	–
Cang Zhu	43.14	–
Chen Pi	43.25	–
Gan Cao	44.01	–
Chuan Xiong	44.32	–
Yan Hu Suo	45.00	–
Fo Shou	45.85	–
Zhe Bei Mu	46.04	–
Dong Gua Zi	46.33	–
She Gan	47.29	–
Gou Qi Zi	47.29	–
Da Zao	47.31	–
Fu Zi	47.60	–
Sheng Jiang	47.60	–
Gou Teng	47.76	–
Ce Bai Ye	48.31	–
Jiao Gu Lan	48.37	–
Qiang Huo	48.50	–
Dan Dou Chi	48.54	–
Wu Wei Zi	48.63	–
Yi Yi Ren	48.77	–
Gui Zhi	49.22	–
Ji Xue Teng	49.40	–
Xiang Fu	49.67	–
Gu Sui Bu	49.75	–
Jiang Huang	49.82	–
Ba Ji Tian	49.98	–
Shen Qu	50.19	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Zhu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62385	62385	0.00	6.54
62386	62386	0.00	5.95
62975	62975	0.00	9.97
62976	62976	0.00	9.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bai Zi Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60197-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bai Zi Ren; Biotae orientalis semen

Special notes

When selecting the *Bai Zi Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bai Zi Ren	1	0	3

Second-stage model

For differentiation of the substance/substance group *Bai Zi Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bai Zi Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Zi Ren	G052HS187SK1	62659	40	from supplier
PhytoComm	Bai Zi Ren	G052HS187SK1	62660	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Bai Zi Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Bai Zi Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Zi Ren	G052HS187SK1	62659†	20
PhytoComm	Bai Zi Ren	G052HS187SK1	62660†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Bai Zi Ren*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Zi Ren	g052f110315	1
Phytocomm	Bai Zi Ren	G052HS187	2
Phytocomm	Bai Zi Ren	g952f110315	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bai Zi Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bai Zi Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	4	853

The substance/substance group *Bai Zi Ren* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bai Zi Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lian Zi	5.91	–
Ji Li	7.73	–
Bai Shao Yao	9.98	–
Ling Zhi	9.99	–
Ze Xie	10.92	–
Shen Qu	10.99	–
Huo Ma Ren	11.23	–
He Huan Pi	11.25	–
Sheng Jiang	11.97	–
Tai Zi Shen	12.46	–
Di Gu Pi	12.71	–
Lai Fu Zi	12.75	–
Zhu Ru	12.81	–
Yuan Zhi	13.70	–
Zhe Bei Mu	14.01	–
Gua Lou	14.40	–
Fu Xiao Mai	14.43	–
Ci Wu Jia	14.64	–
Tian Hua Fen	14.72	–
Gui Zhi	14.74	–
Shan Yao	14.77	–
Pi Pa Ye	15.73	–
Mu Zei	16.29	–
Ren Dong Teng	16.34	–
Tao Ren	16.83	–
Fu Ling	17.19	–
Ye Jiao Teng	17.34	–
Tu Fu Ling	17.62	–
Bai Xian Pi	17.63	–
San Qi	18.04	–
Chuan Xiong	18.09	–
Cang Zhu	18.22	–
Gu Sui Bu	18.67	–
Suan Zao Ren	18.74	–
Yi Yi Ren	18.79	–
Lian Qiao	19.08	–
Chen Pi	19.43	–
Rou Gui	20.21	–
Lu Gen	20.62	–
Gou Teng	20.69	–
Yu Jin	20.69	–
Dang Gui Wei	21.14	–
Fo Shou	22.73	–
Mao Dong Qing	23.29	–
Jin Yin Hua	23.40	–
Fu Zi	23.71	–
Yu Zhu	24.08	–
Ji Xue Teng	24.26	–
Ban Xia (Jiang)	24.31	–
Dan Dou Chi	24.55	–
Guang Huo Xiang	24.63	–
Jie Geng	24.64	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	24.76	–
Ce Bai Ye	24.76	–
She Gan	24.98	–
Fu Pen Zi	25.57	–
Chai Hu	26.08	–
Dan Shen	26.45	–
Hou Po	26.48	–
Ban Lan Gen	26.88	–
Dang Gui	27.17	–
Zhi Mu	27.26	–
Ma Huang Gen	28.05	–
Che Qian Zi	28.11	–
Shan Yu Rou	29.02	–
Zhu Ling	29.34	–
(Fen) Bi Xie	29.51	–
Zhi Ke	30.19	–
Mu Gua	31.21	–
Yan Hu Suo	31.36	–
Ren Shen	32.97	–
Ku Shen	33.19	–
Xie Bai	33.38	–
Hong Jing Tian	33.83	–
Chuang Mu Xiang	33.87	–
Yin Yang Huo	34.21	–
Mang Xiao	34.41	–
Gou Qi Zi	34.53	–
Gan Cao	35.75	–
Wu Wei Zi	36.43	–
Fu Shen	37.31	–
Qiang Huo	37.71	–
Ma Huang	38.81	–
Huang Qin	40.33	–
Long Yan Rou	42.13	–
Zi Hua Di Ding	43.66	–
Zhi Gan Cao	44.84	–
Qing Pi	45.09	–
Jing Jie	45.84	–
Sang Zhi	46.61	–
Huang Lian	47.06	–
Bo He	47.19	–
Du Zhong	47.31	–
Huang Bai	51.18	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Zi Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62659	62659	0.00	6.19
62660	62660	0.00	5.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ban Lan Gen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60230-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ban Lan Gen; Isatidis radix

Special notes

When selecting the *Ban Lan Gen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ban Lan Gen	1	0	1

Second-stage model

For differentiation of the substance/substance group *Ban Lan Gen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ban Lan Gen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ban Lan Gen	G128HS165TL1	62867	40	from supplier
PhytoComm	Ban Lan Gen	G128HS165TL1	62868	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ban Lan Gen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ban Lan Gen*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ban Lan Gen	G128HS165TL1	62867†	20
PhytoComm	Ban Lan Gen	G128HS165TL1	62868†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Ban Lan Gen*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ban Lan Gen	G128H0816421	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ban Lan Gen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ban Lan Gen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	2	855

The substance/substance group *Ban Lan Gen* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ban Lan Gen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	6.90	–
Ze Xie	8.43	–
Gua Lou	9.77	–
Tian Hua Fen	10.03	–
Fu Zi	10.64	–
Hou Po	12.28	–
Hong Jing Tian	12.33	–
Mu Zei	12.33	–
Ren Dong Teng	12.41	–
Pi Pa Ye	12.65	–
Cang Zhu	12.77	–
Dan Dou Chi	12.90	–
Chen Pi	13.04	–
Ma Huang	13.22	–
Gou Qi Zi	13.90	–
Lian Qiao	14.15	–
Ku Shen	14.23	–
Shan Yao	14.33	–
Fu Pen Zi	14.37	–
Dang Gui	14.60	–
Bai Shao Yao	14.82	–
Chuan Xiong	14.84	–
Yu Zhu	15.05	–
Guang Huo Xiang	15.13	–
Jiao Gu Lan	15.26	–
Ren Shen	15.55	–
Ji Xue Teng	16.04	–
Yuan Zhi	16.56	–
Lu Gen	16.58	–
Yan Hu Suo	16.67	–
Zhe Bei Mu	16.73	–
Dang Gui Wei	17.46	–
Jie Geng	17.54	–
Tai Zi Shen	17.70	–
He Huan Pi	18.04	–
Lai Fu Zi	18.38	–
Ye Jiao Teng	18.64	–
Gu Sui Bu	18.70	–
Zhi Ke	18.77	–
Gui Zhi	18.97	–
Ji Li	19.07	–
Zhi Gan Cao	19.16	–
Tao Ren	19.29	–
Dan Shen	19.36	–
Ling Zhi	19.71	–
Yin Yang Huo	19.86	–
Zi Hua Di Ding	20.03	–
San Qi	20.18	–
Shen Qu	20.33	–
Wu Wei Zi	20.35	–
Suan Zao Ren	20.86	–
Gan Cao	21.08	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Mu	21.31	–
Che Qian Zi	21.32	–
Tu Fu Ling	21.40	–
Huo Ma Ren	21.44	–
Qiang Huo	21.58	–
Bai Zi Ren	21.69	–
Rou Gui	21.74	–
Huang Qin	21.92	–
Chai Hu	22.60	–
Long Yan Rou	22.82	–
Gou Teng	22.89	–
Ce Bai Ye	23.45	–
Lian Zi	23.53	–
Ci Wu Jia	23.68	–
Chuang Mu Xiang	23.69	–
Sheng Jiang	23.89	–
Di Gu Pi	24.09	–
She Gan	24.38	–
Fu Ling	24.62	–
Xie Bai	24.69	–
Qing Pi	24.80	–
(Fen) Bi Xie	24.82	–
Mao Dong Qing	24.95	–
Bai Xian Pi	25.05	–
Shan Yu Rou	25.31	–
Yu Jin	26.17	–
Ban Xia (Jiang)	26.55	–
Hong Hua	27.22	–
Zhu Ling	27.32	–
Zhu Ru	28.62	–
Mu Gua	28.81	–
Huang Lian	29.07	–
Bo He	29.43	–
Fo Shou	29.87	–
Mang Xiao	30.79	–
Sha Ren	31.30	–
Jing Jie	32.18	–
Chi Shao (Yao)	32.55	–
Du Zhong	33.47	–
Sang Zhi	33.63	–
Yi Yi Ren	33.90	–
Huang Bai	34.11	–
Fu Xiao Mai	35.00	–
Ma Huang Gen	35.58	–
Sang Ye	38.45	–
Cang Er Zi	39.32	–
Fu Shen	42.13	–
Bai Jiang Cao	44.30	–
Qing Hao	48.21	–
Huang Qi	50.25	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ban Lan Gen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62867	62867	0.00	7.94
62868	62868	0.00	6.90

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ban Xia (Jiang)**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60019-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ban Xia (Jiang); Pinelliae rhizoma praeparatum cum zingibere

Special notes

When selecting the *Ban Xia (Jiang)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ban Xia (Jiang)	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ban Xia (Jiang)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ban Xia (Jiang)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62599	40	from supplier
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62600	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ban Xia (Jiang)*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ban Xia (Jiang)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62599†	20
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62600†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Ban Xia (Jiang)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ban Xia (Jiang)	g191h0526121	1
Phytocomm	Ban Xia (Jiang)	G191H0526322	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ban Xia (Jiang)* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ban Xia (Jiang)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	79	1	24 520
Type B	10	39	1	12 427
Type C	0	0	2	855

The substance/substance group *Ban Xia (Jiang)* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	98.7500 % (> 95.0000 %)
Type B	99.9713 % (> 99.9416 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ban Xia (Jiang)* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Shao Yao	4.24	–
Di Gu Pi	5.87	–
Gui Zhi	7.20	–
Sheng Jiang	7.59	–
Rou Gui	8.41	–
Zhu Ru	8.87	–
He Huan Pi	9.32	–
Fu Xiao Mai	9.37	–
Shen Qu	11.10	–
Ci Wu Jia	11.57	–
Zhe Bei Mu	11.76	–
Ji Xue Teng	12.02	–
Tao Ren	12.35	–
Yi Yi Ren	12.56	–
Tai Zi Shen	12.83	–
Gou Teng	13.68	–
Fu Ling	14.23	–
Tian Hua Fen	14.34	–
Fo Shou	14.54	–
Bai Zi Ren	14.74	–
Ji Li	15.35	–
Ye Jiao Teng	15.42	–
Lian Zi	16.72	–
Mu Zei	16.76	–
Huo Ma Ren	17.29	–
Gua Lou	17.99	–
Bai Xian Pi	18.67	–
Ling Zhi	18.94	–
Lai Fu Zi	19.27	–
Ma Huang Gen	19.80	–
Pi Pa Ye	20.14	–
Fu Shen	20.22	–
Gu Sui Bu	20.67	–
Lu Gen	21.12	–
Yuan Zhi	21.17	–
Lian Qiao	22.31	–
Yan Hu Suo	24.93	–
Jin Yin Hua	25.53	–
Shan Yao	25.88	–
Zhu Ling	26.94	–
Ren Dong Teng	27.39	–
Yu Jin	27.49	–
Chen Pi	27.49	–
Chuan Xiong	27.59	–
Tu Fu Ling	27.72	–
Ze Xie	27.98	–
Ren Shen	28.86	–
Zhi Mu	29.38	–
(Fen) Bi Xie	29.96	–
Dang Gui	30.34	–
Huang Qin	30.75	–
Fu Zi	31.07	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jie Geng	31.41	–
Shan Yu Rou	31.57	–
Jiao Gu Lan	31.74	–
Zhi Ke	32.17	–
Fu Pen Zi	32.33	–
Dang Gui Wei	33.00	–
Mang Xiao	33.06	–
Ce Bai Ye	33.55	–
Suan Zao Ren	33.65	–
Dan Shen	33.85	–
Yu Zhu	33.87	–
Dan Dou Chi	34.01	–
Cang Zhu	34.27	–
Yin Yang Huo	35.33	–
Ban Lan Gen	36.37	–
Hou Po	37.47	–
San Qi	37.73	–
She Gan	38.38	–
Ma Huang	38.51	–
Wu Wei Zi	38.85	–
Gou Qi Zi	40.23	–
Chai Hu	40.27	–
Ku Shen	40.70	–
Mao Dong Qing	41.28	–
Mu Gua	41.66	–
Guang Huo Xiang	44.37	–
Che Qian Zi	44.46	–
Qiang Huo	45.49	–
Long Yan Rou	46.58	–
Gan Cao	46.88	–
Huang Lian	47.46	–
Sang Zhi	47.63	–
Hong Jing Tian	48.22	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ban Xia (Jiang)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62599	62599	0.00	4.90
62600	62600	0.00	4.24

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all

substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ban Zhi Lian**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60038-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ban Zhi Lian; Scutellariae barbatae herba

Special notes

When selecting the *Ban Zhi Lian* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ban Zhi Lian	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ban Zhi Lian* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ban Zhi Lian*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ban Zhi Lian	G224H0528821	62753	40	from supplier
PhytoComm	Ban Zhi Lian	G224H0528821	62754	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ban Zhi Lian*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ban Zhi Lian*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ban Zhi Lian	G224H0528821	62753†	20
PhytoComm	Ban Zhi Lian	G224H0528821	62754†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Ban Zhi Lian*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ban Zhi Lian	G224H0528221	1
Phytocomm	Ban Zhi Lian	G224H0528521	2
PhytoComm	Ban Zhi Lian	G224H0528521	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ban Zhi Lian* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ban Zhi Lian* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	39	1	12 437
Type C	0	0	4	853

The substance/substance group *Ban Zhi Lian* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ban Zhi Lian* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mai Ya	22.59	–
Sang Zhi	23.14	–
Fu Ling	25.28	–
Mang Xiao	25.76	–
Chuan Mu Tong	25.79	–
Rou Gui	26.12	–
Bai He	26.56	–
E Zhu	27.69	–
Bai Xian Pi	28.72	–
Tian Hua Fen	28.83	–
Yan Hu Suo	30.53	–
Ji Xue Teng	32.46	–
Sha Shen (Bei)	32.50	–
Shan Yao	32.81	–
Jiang Huang	32.88	–
Ji Li	33.15	–
Di Gu Pi	33.18	–
Chuan Lian Zi	33.34	–
Mi Huan Jun	33.67	–
Lian Zi	34.01	–
Ye Jiao Teng	35.15	–
Ling Zhi	35.16	–
Gou Teng	35.17	–
Dang Gui	35.43	–
Bai Shao Yao	35.57	–
Long Dan (Cao)	35.70	–
Chai Hu	35.80	–
Gan Jiang	36.11	–
Yi Yi Ren	36.11	–
Dan Dou Chi	37.95	–
He Huan Pi	38.52	–
Fo Shou	38.88	–
(Bai) Dou Kou	39.22	–
Du Zhong	39.26	–
Mao Dong Qing	39.35	–
Tu Fu Ling	39.50	–
Jie Geng	39.59	–
Zhi Ke	39.68	–
Shen Qu	40.11	–
Zi Hua Di Ding	40.56	–
Ren Dong Teng	40.88	–
Wu Wei Zi	41.04	–
Ma Huang Gen	41.17	–
Bo He	41.26	–
Lai Fu Zi	41.28	–
(Shi) Chang Pu	41.40	–
Yu Jin	42.47	–
Qin Jiao	42.77	–
Sheng Jiang	42.78	–
Zhu Ling	43.27	–
Da Zao	43.81	–
Qiang Huo	43.91	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Zhu	43.95	–
Huang Lian	44.09	–
Sha Ren	44.71	–
Huang Qi	44.74	–
Zhu Ru	46.53	–
Cang Er Zi	46.70	–
Jin Yin Hua	46.72	–
Yuan Zhi	47.07	–
Gui Zhi	48.49	–
Tai Zi Shen	48.56	–
Zi Su Zi	48.91	–
Fu Shen	49.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ban Zhi Lian* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62753	62753	0.00	22.85
62754	62754	0.00	22.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bing Lang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60186-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bing Lang; Arecae catechu semen

Special notes

When selecting the *Bing Lang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bing Lang	1	0	0

Second-stage model

For differentiation of the substance/substance group *Bing Lang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bing Lang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bing Lang	G027H1825021	63009	40	from supplier
PhytoComm	Bing Lang	G027H1825021	63010	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Bing Lang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Bing Lang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bing Lang	G027H1825021	63009†	20
PhytoComm	Bing Lang	G027H1825021	63010†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Bing Lang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bing Lang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bing Lang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	0	857

The substance/substance group *Bing Lang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bing Lang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Huang Qi	22.47	–
Mang Xiao	25.72	–
Dang Gui	26.68	–
Sang Zhi	31.25	–
Bai Zhi	31.30	–
Chuan Niu Xi	32.49	–
(Huai) Niu Xi	33.50	–
(Shi) Chang Pu	34.78	–
Qin Jiao	35.53	–
Da Zao	36.71	–
Jie Geng	37.84	–
Sha Shen (Bei)	38.62	–
Chen Pi	39.13	–
Tian Hua Fen	39.20	–
Long Dan (Cao)	39.47	–
Chi Shao (Yao)	39.94	–
Mu Dan Pi	40.86	–
Ju Hua	40.94	–
Bai Zhu	41.12	–
Shan Yu Rou	41.78	–
Du Huo	42.49	–
Zi Su Zi	44.02	–
Zhi Gan Cao	44.05	–
Fang Feng	44.54	–
He Huan Pi	44.62	–
Di Gu Pi	45.48	–
Chuan Lian Zi	45.59	–
Mu Gua	46.89	–
Lian Qiao	47.39	–
E Zhu	48.38	–
Shan Yao	48.60	–
Sha Ren	49.08	–
Zhi Ke	49.23	–
Mi Huan Jun	49.25	–
Gu Sui Bu	49.96	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bing Lang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
63009	63009	0.00	22.79
63010	63010	0.00	22.47

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by

laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bo He**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60073-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bo He; Menthae haplocalysis herba

Special notes

When selecting the *Bo He* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bo He	1	0	2

Second-stage model

For differentiation of the substance/substance group *Bo He* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bo He*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bo He	G158HS332TH1	62863	40	from supplier
PhytoComm	Bo He	G158HS332TH1	62864	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Bo He*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Bo He*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bo He	G158HS332TH1	62863†	20
PhytoComm	Bo He	G158HS332TH1	62864†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Bo He*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bo He	g158h1712022	1
PhytoComm	Bo He	G158H1712121	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bo He* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bo He* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	34	6	12 437
Type C	0	0	2	855

The substance/substance group *Bo He* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	85.0000 % (> 77.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bo He* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chai Hu	4.96	–
Sha Ren	6.08	–
Fu Zi	8.15	–
Du Zhong	9.20	–
Zi Hua Di Ding	9.27	–
Dan Shen	9.50	–
Qiang Huo	9.78	–
Bai Jiang Cao	10.79	–
Yu Jin	10.91	–
Jing Jie	11.05	–
Mao Dong Qing	11.17	–
Qing Hao	11.20	–
Hou Po	12.04	–
Jiao Gu Lan	12.14	–
Huang Bai	12.24	–
Yin Yang Huo	14.08	–
Ce Bai Ye	14.12	–
Bai Xian Pi	14.21	–
Shan Yao	14.75	–
Dan Dou Chi	14.83	–
Fu Pen Zi	15.00	–
Ye Jiao Teng	15.83	–
Huang Lian	15.84	–
Jin Yin Hua	16.38	–
Che Qian Zi	16.64	–
Qing Pi	17.12	–
Zhi Ke	17.39	–
Pi Pa Ye	17.69	–
Yan Hu Suo	17.70	–
Hong Jing Tian	17.95	–
Chuan Xiong	18.11	–
Ji Xue Teng	18.28	–
Ling Zhi	18.62	–
Pu Gong Ying	18.90	–
Tian Hua Fen	19.52	–
Shen Qu	19.75	–
Ma Huang	19.93	–
Tu Fu Ling	20.11	–
Gu Sui Bu	20.16	–
Guang Huo Xiang	20.70	–
Lian Zi	20.92	–
Ren Dong Teng	20.99	–
Wu Wei Zi	21.48	–
Di Gu Pi	22.38	–
Bai Shao Yao	22.38	–
Gan Cao	23.04	–
Ji Li	23.92	–
Ze Lan	24.68	–
Suan Zao Ren	24.70	–
(Fen) Bi Xie	25.80	–
Gua Lou	25.87	–
Ban Lan Gen	26.36	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhu Ling	26.53	–
He Huan Pi	26.54	–
Ren Shen	26.90	–
Yuan Zhi	27.81	–
Sang Ye	27.88	–
Lian Qiao	28.48	–
She Gan	28.54	–
Hong Hua	28.58	–
Gou Qi Zi	28.73	–
Chuang Mu Xiang	28.85	–
Gou Teng	29.12	–
Zhi Gan Cao	29.53	–
Lu Gen	29.88	–
Mang Xiao	30.48	–
Mu Zei	30.64	–
Cang Zhu	30.76	–
Xi Xian Cao	30.93	–
Ma Huang Gen	31.86	–
Chi Shao (Yao)	32.11	–
Rou Gui	32.19	–
Zhe Bei Mu	33.63	–
Cang Er Zi	34.25	–
Sheng Jiang	34.50	–
Huo Ma Ren	35.40	–
Ban Zhi Lian	36.64	–
Jie Geng	36.67	–
Nü Zhen Zi	37.80	–
Tao Ren	38.03	–
Huang Qin	38.18	–
Sang Zhi	38.38	–
Fu Ling	38.56	–
Shan Yu Rou	39.47	–
E Zhu	39.67	–
Jiang Huang	40.02	–
Gui Zhi	40.52	–
Chen Pi	40.63	–
Zhi Shi	40.86	–
Yi Yi Ren	41.52	–
Xie Bai	42.52	–
Wu Mei	42.80	–
(Shi) Chang Pu	43.69	–
Tai Zi Shen	43.73	–
Yi Mu Cao	43.89	–
Sha Shen (Bei)	44.20	–
Zi Su Zi	44.58	–
Ze Xie	44.69	–
Dang Gui	44.83	–
Ci Wu Jia	44.90	–
Ban Xia (Jiang)	45.47	–
Long Yan Rou	45.87	–
Xin Yi	46.22	–
Niu Bang Zi	46.60	–
Bai Zi Ren	47.18	–
Yin Chen Hao	47.19	–
Ku Shen	47.24	–
San Qi	48.10	–
Yu Zhu	48.52	–
Fo Shou	48.55	–
Fu Shen	48.76	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Qi	49.44	–
Zhu Ru	49.55	–
Dang Gui Wei	49.82	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bo He* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62863	62863	0.00	5.92
62864	62864	0.00	4.96

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Bu Gu Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60023-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Bu Gu Zhi; Psoraleae corylifoliae fructus

Special notes

When selecting the *Bu Gu Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Bu Gu Zhi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Bu Gu Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Bu Gu Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bu Gu Zhi	G205H1247822	62487	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247822	62488	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247922	62851	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247922	62852	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Bu Gu Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Bu Gu Zhi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bu Gu Zhi	G205H1247822	62487†	20
PhytoComm	Bu Gu Zhi	G205H1247822	62488†	20
PhytoComm	Bu Gu Zhi	G205H1247922	62851†	20
PhytoComm	Bu Gu Zhi	G205H1247922	62852†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Bu Gu Zhi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Bu Gu Zhi	g205h1247222	1
Phytocomm	Bu Gu Zhi	G205H1247222	2
Phytocomm	Bu Gu Zhi	G205H1247421	2

- 852 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Bu Gu Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Bu Gu Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	5	852

The substance/substance group *Bu Gu Zhi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Bu Gu Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Tu Si Zi	11.66	–
Dan Zhu Ye	18.67	–
Gu Sui Bu	21.45	–
Jing Jie	21.59	–
Sha Ren	22.80	–
Yu Xing Cao	23.85	–
Wu Yao	25.65	–
Yin Chen Hao	25.78	–
Mang Xiao	26.45	–
Sang Bai Pi	26.58	–
Sang Ji Shend	28.19	–
Wu Zhu Yu	28.80	–
He Huan Pi	29.73	–
He Shou Wu	30.23	–
Ji Li	31.86	–
Xuan Fu Hua	33.00	–
Wu Jia Pi	33.38	–
(Bai) Dou Kou	34.48	–
Chen Pi	35.62	–
Xian Mao	36.41	–
Jiang Huang	37.98	–
Jin Qian Cao	38.68	–
Ge Gen	39.33	–
(Sheng) Di Huang	43.01	–
(Shi) Chang Pu	44.29	–
Shan Yao	45.00	–
Du Zhong	46.49	–
E Zhu	47.18	–
Nü Zhen Zi	48.03	–
Hua Shi	48.19	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bu Gu Zhi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62487	62487	0.00	18.78
62488	62488	0.00	18.67
62851	62851	0.00	11.66
62852	62852	0.00	12.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Cang Er Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60054-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Cang Er Zi; Xanthii fructus

Special notes

When selecting the *Cang Er Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Cang Er Zi	2	0	3

Second-stage model

For differentiation of the substance/substance group *Cang Er Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Cang Er Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Cang Er Zi	G251H1401821	62503	40	from supplier
PhytoComm	Cang Er Zi	G251H1401821	62504	40	from supplier
PhytoComm	Cang Er Zi	G251H1401921	62755	40	from supplier
PhytoComm	Cang Er Zi	G251H1401921	62756	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Cang Er Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Cang Er Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Cang Er Zi	G251H1401821	62503†	20
PhytoComm	Cang Er Zi	G251H1401821	62504†	20
PhytoComm	Cang Er Zi	G251H1401921	62755†	20
PhytoComm	Cang Er Zi	G251H1401921	62756†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Cang Er Zi*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Cang Er Zi	g251h1401021	1
Phytocomm	Cang Er Zi	G251H1401322	1
PhytoComm	Cang Er Zi	G251H1401322	1
Phytocomm	Cang Er Zi	G251H1401521	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Cang Er Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Cang Er Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	1	0	4	852

The substance/substance group *Cang Er Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.5349 % (> 98.9483 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Cang Er Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zi Su Zi	10.30	–
Shan Yao	11.27	–
Gan Jiang	14.77	–
E Zhu	16.39	–
Ji Li	17.26	–
Suan Zao Ren	18.05	–
Wu Wei Zi	20.30	–
Wu Mei	20.89	–
Sha Shen (Bei)	21.11	–
Niu Bang Zi	21.33	–
Sang Zhi	22.10	–
Zi Hua Di Ding	23.45	–
Jiang Huang	24.54	–
Jin Yin Hua	24.56	–
Di Gu Pi	25.34	–
Yan Hu Suo	25.52	–
Mang Xiao	26.19	–
Chuan Lian Zi	26.20	–
Qing Pi	27.88	–
Yi Mu Cao	28.22	–
Lai Fu Zi	29.22	–
Jie Geng	29.72	–
(Huai) Niu Xi	29.88	–
Tian Hua Fen	30.09	–
Shan Yu Rou	30.13	–
Sha Ren	30.25	–
Zhi Ke	30.42	–
Ku Shen	30.57	–
Chen Pi	30.67	–
Qiang Huo	30.83	–
Jiao Gu Lan	31.16	–
Gua Lou	31.71	–
(Bai) Dou Kou	32.14	–
Sang Ye	32.23	–
Du Zhong	32.84	–
Lian Qiao	32.93	–
Huang Bai	33.06	–
Hou Po	33.44	–
Xiang Fu	34.03	–
Ze Lan	34.06	–
Gan Cao	34.53	–
Huang Lian	34.68	–
Mao Dong Qing	34.88	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pu Gong Ying	34.91	–
Long Dan (Cao)	35.99	–
Ce Bai Ye	36.12	–
Dan Dou Chi	36.29	–
Dang Gui	36.37	–
Fu Zi	36.48	–
(Shi) Chang Pu	37.21	–
Shen Qu	37.23	–
Hong Hua	37.38	–
He Huan Pi	37.83	–
Xi Xian Cao	38.04	–
Chi Shao (Yao)	38.27	–
Dong Gua Zi	38.62	–
Chuan Xiong	38.68	–
Chuan Mu Tong	39.26	–
Huang Qi	39.54	–
Dan Shen	39.65	–
Tu Fu Ling	40.28	–
Bo He	40.43	–
Mi Huan Jun	40.72	–
Nü Zhen Zi	40.86	–
Xiao Hui Xiang	41.26	–
Zhe Bei Mu	42.01	–
Qing Hao	42.38	–
Ren Dong Teng	42.54	–
Mai Ya	42.62	–
Xin Yi	42.67	–
Hong Jing Tian	42.69	–
Yu Jin	42.82	–
Jing Jie	43.18	–
Bai Zhu	43.26	–
Ye Jiao Teng	43.30	–
Xie Bai	43.38	–
Cang Zhu	43.40	–
Bai Xian Pi	43.95	–
Fu Ling	43.99	–
Lian Zi	43.99	–
Gou Qi Zi	44.45	–
(Fen) Bi Xie	44.48	–
Yuan Zhi	44.72	–
Chai Hu	44.96	–
Du Huo	45.01	–
Chuan Niu Xi	45.15	–
Bai Hua She She Cao	45.22	–
Wang Bu Liu Xing	45.24	–
Yin Yang Huo	45.43	–
Bai Jiang Cao	45.60	–
Zhi Shi	45.70	–
Chuang Mu Xiang	46.23	–
Da Zao	46.51	–
Che Qian Zi	46.62	–
Bai He	46.87	–
Ling Zhi	47.08	–
Zhi Gan Cao	47.80	–
Fu Pen Zi	47.97	–
Ban Zhi Lian	48.85	–
Ban Lan Gen	48.90	–
Ma Huang	49.43	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Cang Er Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62503	62503	0.00	11.27
62504	62504	0.00	11.29
62755	62755	0.00	10.30
62756	62756	0.00	10.45

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Cang Zhu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60160-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Cang Zhu; Atractylodis rhizoma

Special notes

When selecting the *Cang Zhu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Cang Zhu	2	0	1

Second-stage model

For differentiation of the substance/substance group *Cang Zhu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Cang Zhu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Cang Zhu	G042HS301SG1	62653	40	from supplier
PhytoComm	Cang Zhu	G042HS301SG1	62654	40	from supplier
PhytoComm	Cang Zhu	G042HS301TG1	62977	40	from supplier
PhytoComm	Cang Zhu	G042HS301TG1	62978	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Cang Zhu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Cang Zhu*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Cang Zhu	G042HS301SG1	62653†	20
PhytoComm	Cang Zhu	G042HS301SG1	62654†	20
PhytoComm	Cang Zhu	G042HS301TG1	62977†	20
PhytoComm	Cang Zhu	G042HS301TG1	62978†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Cang Zhu*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Cang Zhu	g042h1402221	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Cang Zhu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Cang Zhu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	160	0	24 439
Type B	7	77	3	12 390
Type C	0	0	1	856

The substance/substance group *Cang Zhu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	100.0000 % (> 96.2500 %)
Type B	99.9321 % (> 99.9023 %)	96.2500 % (> 92.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Cang Zhu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yu Zhu	5.38	–
Dang Gui	6.00	–
Dang Gui Wei	6.35	–
Shan Yao	7.73	–
Xie Bai	8.36	–
Jie Geng	9.00	–
Pi Pa Ye	9.08	–
Chen Pi	9.24	–
Yuan Zhi	9.68	–
Gua Lou	10.15	–
Lian Qiao	10.16	–
Ku Shen	10.26	–
Gou Qi Zi	10.51	–
Chuan Xiong	10.91	–
Suan Zao Ren	11.61	–
Jiao Gu Lan	12.31	–
Chuang Mu Xiang	12.65	–
Bai Shao Yao	12.73	–
San Qi	12.94	–
Ji Li	13.14	–
Jin Yin Hua	13.67	–
Tian Hua Fen	13.80	–
Tai Zi Shen	14.07	–
Bai Zi Ren	14.56	–
Shen Qu	14.60	–
Zhe Bei Mu	14.88	–
Lian Zi	14.92	–
Ze Xie	15.15	–
Mu Gua	15.19	–
Ren Shen	15.42	–
Zhi Mu	15.51	–
Ling Zhi	15.78	–
Long Yan Rou	15.79	–
Mu Zei	16.30	–
Mao Dong Qing	16.72	–
Dan Dou Chi	16.93	–
Ren Dong Teng	17.28	–
Dan Shen	17.59	–
Ban Lan Gen	17.65	–
Di Gu Pi	17.66	–
Ye Jiao Teng	17.86	–
Lai Fu Zi	18.27	–
Hou Po	19.17	–
Shan Yu Rou	19.40	–
He Huan Pi	19.97	–
Zhi Ke	20.08	–
Hong Jing Tian	20.31	–
Fo Shou	20.60	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huo Ma Ren	21.12	–
Fu Ling	21.25	–
Zhu Ru	21.25	–
Chai Hu	21.47	–
Guang Huo Xiang	21.77	–
Fu Zi	21.82	–
Sheng Jiang	21.93	–
Gu Sui Bu	22.51	–
Wu Wei Zi	22.67	–
Tao Ren	23.00	–
Ci Wu Jia	23.04	–
Zhi Gan Cao	23.83	–
Fu Pen Zi	23.98	–
Ma Huang	24.07	–
Lu Gen	24.35	–
Tu Fu Ling	24.64	–
Gou Teng	24.98	–
Rou Gui	25.61	–
Gui Zhi	26.21	–
Gan Cao	26.25	–
Ji Xue Teng	26.93	–
Bai Xian Pi	27.41	–
She Gan	27.72	–
Huang Qin	27.92	–
Yan Hu Suo	29.80	–
Che Qian Zi	29.81	–
Qing Pi	30.57	–
Yin Yang Huo	30.61	–
Zi Hua Di Ding	30.75	–
Qiang Huo	30.83	–
Yu Jin	30.91	–
Fu Xiao Mai	30.94	–
Mang Xiao	31.25	–
Ban Xia (Jiang)	31.45	–
Ce Bai Ye	32.02	–
Zhu Ling	32.98	–
Hong Hua	33.05	–
Cang Er Zi	33.52	–
Yi Yi Ren	33.65	–
Bo He	34.51	–
Mai Men Dong	35.91	–
Ma Huang Gen	36.47	–
(Fen) Bi Xie	36.76	–
Sang Zhi	36.77	–
Sha Ren	38.94	–
Chi Shao (Yao)	39.07	–
Huang Bai	39.21	–
Du Zhong	39.41	–
Huang Lian	41.23	–
Jing Jie	42.34	–
Fu Shen	43.64	–
Chuan Lian Zi	48.76	–
Huang Qi	49.15	–
Qing Hao	49.75	–
Sang Ye	49.79	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Cang Zhu* is separated from critical neighbours in a second-stage model, all

the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62653	62653	0.00	7.73
62654	62654	0.00	7.94
62977	62977	0.00	6.20
62978	62978	0.00	5.38

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ce Bai Ye**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60313-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ce Bai Ye; Biotae orientalis cacumen

Special notes

When selecting the *Ce Bai Ye* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ce Bai Ye	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ce Bai Ye* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ce Bai Ye*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ce Bai Ye	G050HS245TG1	62925	40	from supplier
PhytoComm	Ce Bai Ye	G050HS245TG1	62926	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ce Bai Ye*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ce Bai Ye*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ce Bai Ye	G050HS245TG1	62925†	20
PhytoComm	Ce Bai Ye	G050HS245TG1	62926†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 2 *Apo-Ident* customers from 3 batches from the substance/substance group *Ce Bai Ye*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
phytocomm	Ce Bai Ye	g050h0975321	2
Phytocomm	Ce Bai Ye	G050H0975321	1
Phytocomm	Ce Bai Ye	G050HS245	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ce Bai Ye* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ce Bai Ye* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	6	79	1	24 514
Type B	12	36	4	12 425
Type C	0	0	4	853

The substance/substance group *Ce Bai Ye* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9786 % (> 99.9636 %)	98.7500 % (> 95.0000 %)
Type B	99.9143 % (> 99.8846 %)	90.0000 % (> 82.5000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ce Bai Ye* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dan Dou Chi	2.60	–
Guang Huo Xiang	5.21	–
Bai Xian Pi	5.27	–
Ling Zhi	6.20	–
Ren Dong Teng	7.46	–
Mao Dong Qing	8.00	–
Lian Zi	8.07	–
Tu Fu Ling	8.53	–
Chai Hu	8.76	–
Shen Qu	8.91	–
Zhu Ling	10.33	–
Yu Jin	10.61	–
Yin Yang Huo	10.68	–
Hou Po	10.81	–
(Fen) Bi Xie	11.29	–
Ji Li	11.47	–
Jing Jie	11.86	–
Ye Jiao Teng	12.25	–
Gu Sui Bu	12.29	–
Suan Zao Ren	12.37	–
She Gan	12.61	–
Che Qian Zi	12.73	–
Qiang Huo	13.56	–
Shan Yao	13.78	–
Dan Shen	14.01	–
Chuan Xiong	14.85	–
He Huan Pi	15.04	–
Bai Shao Yao	15.04	–
Fu Pen Zi	15.07	–
Huo Ma Ren	15.98	–
Rou Gui	16.07	–
Zi Hua Di Ding	16.17	–
Pi Pa Ye	16.31	–
Fu Ling	16.48	–
Lian Qiao	16.48	–
Sheng Jiang	16.51	–
Gou Teng	16.87	–
Fu Zi	16.90	–
Lu Gen	17.16	–
Gan Cao	17.41	–
Gua Lou	17.91	–
Gui Zhi	17.95	–
Yan Hu Suo	18.04	–
Jin Yin Hua	18.23	–
Ji Xue Teng	19.71	–
Tian Hua Fen	19.89	–
Jiao Gu Lan	19.94	–
Ma Huang Gen	20.17	–
Sha Ren	20.29	–
Ban Lan Gen	20.36	–
Hong Jing Tian	20.66	–
Zhe Bei Mu	20.95	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Yi Ren	21.09	–
Bo He	21.29	–
Mu Zei	21.56	–
Tao Ren	22.16	–
Di Gu Pi	22.25	–
Cang Zhu	22.67	–
Ma Huang	22.86	–
Chen Pi	24.53	–
Zhi Gan Cao	24.59	–
Yuan Zhi	24.72	–
Ci Wu Jia	24.91	–
Fo Shou	25.00	–
Qing Pi	25.29	–
Du Zhong	25.32	–
Huang Bai	25.45	–
Fu Xiao Mai	25.51	–
Bai Zi Ren	26.05	–
Tai Zi Shen	26.17	–
Chuang Mu Xiang	26.47	–
Zhu Ru	27.25	–
Zhi Ke	27.27	–
Wu Wei Zi	27.65	–
Ban Xia (Jiang)	27.87	–
Gou Qi Zi	28.28	–
Yu Zhu	28.73	–
Ze Xie	28.79	–
Huang Lian	29.32	–
Ren Shen	29.59	–
Dang Gui Wei	31.39	–
Mang Xiao	32.57	–
Sang Zhi	33.51	–
San Qi	34.74	–
Dang Gui	34.92	–
Bai Jiang Cao	35.71	–
Xie Bai	36.10	–
Fu Shen	36.38	–
Jie Geng	36.63	–
Lai Fu Zi	36.85	–
Chi Shao (Yao)	37.49	–
Long Yan Rou	38.22	–
Qing Hao	38.37	–
Huang Qin	38.80	–
Hong Hua	42.16	–
Shan Yu Rou	43.34	–
Bai Zhu	44.39	–
Ban Zhi Lian	44.76	–
Cang Er Zi	45.62	–
Ku Shen	46.32	–
Pu Gong Ying	47.22	–
Zhi Mu	49.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ce Bai Ye* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62925	62925	0.00	2.60
62926	62926	0.00	2.66

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chai Hu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60010-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chai Hu; Bupleuri radix

Special notes

When selecting the *Chai Hu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chai Hu	2	0	3

Second-stage model

For differentiation of the substance/substance group *Chai Hu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chai Hu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chai Hu	G055HS233SK1	62661	40	from supplier
PhytoComm	Chai Hu	G055HS233SK1	62662	40	from supplier
PhytoComm	Chai Hu	G055HS233TG1	62905	40	from supplier
PhytoComm	Chai Hu	G055HS233TG1	62906	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Chai Hu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Chai Hu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chai Hu	G055HS233SK1	62661†	20
PhytoComm	Chai Hu	G055HS233SK1	62662†	20
PhytoComm	Chai Hu	G055HS233TG1	62905†	20
PhytoComm	Chai Hu	G055HS233TG1	62906†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Chai Hu*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chai Hu	g055h1055123	1
Phytocomm	Chai Hu	G055H1055123	1
PhytoComm	Chai Hu	G055H1055322	1
Phytocomm	Chai Hu	G055H1055422	5

- 849 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chai Hu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chai Hu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	159	1	24 439
Type B	5	76	4	12 392
Type C	0	0	8	849

The substance/substance group *Chai Hu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9976 % (> 99.9826 %)	99.3750 % (> 97.5000 %)
Type B	99.9286 % (> 99.8987 %)	95.0000 % (> 91.2500 %)
Type C	100.0000 % (> 98.8253 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chai Hu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jing Jie	4.81	–
Ji Li	5.69	–
Bo He	6.07	–
Mao Dong Qing	6.36	–
Qiang Huo	6.62	–
Zi Hua Di Ding	7.05	–
Sha Ren	7.25	–
Ling Zhi	7.48	–
Hou Po	7.67	–
Yin Yang Huo	7.71	–
Yu Jin	7.73	–
Bai Xian Pi	8.11	–
Shen Qu	8.40	–
Fu Zi	8.79	–
Dan Dou Chi	8.84	–
Ce Bai Ye	9.24	–
Lian Zi	9.78	–
Du Zhong	9.91	–
Guang Huo Xiang	9.99	–
Lian Qiao	11.51	–
Shan Yao	11.72	–
Tu Fu Ling	11.76	–
Ye Jiao Teng	12.00	–
Che Qian Zi	12.04	–
Ren Dong Teng	12.72	–
Dan Shen	12.93	–
Chuan Xiong	13.20	–
Huang Bai	13.21	–
Fu Pen Zi	13.35	–
Pi Pa Ye	13.65	–
Yan Hu Suo	14.28	–
Gu Sui Bu	14.42	–
Qing Hao	14.95	–
Jiao Gu Lan	15.28	–
Tian Hua Fen	15.32	–
Bai Shao Yao	15.48	–
Zhu Ling	15.67	–
Suan Zao Ren	15.78	–
Fu Ling	15.92	–
Qing Pi	16.19	–
Bai Jiang Cao	16.25	–
He Huan Pi	16.56	–
Hong Jing Tian	16.76	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	16.89	–
She Gan	16.90	–
Sheng Jiang	17.69	–
Ma Huang	17.91	–
Ji Xue Teng	17.94	–
Ban Lan Gen	17.97	–
Gua Lou	18.47	–
Cang Zhu	18.51	–
(Fen) Bi Xie	18.57	–
Gan Cao	18.88	–
Zhe Bei Mu	19.04	–
Huang Lian	19.71	–
Gui Zhi	20.93	–
Huo Ma Ren	21.27	–
Chen Pi	21.51	–
Zhi Ke	21.71	–
Yuan Zhi	21.93	–
Gou Teng	22.12	–
Rou Gui	22.15	–
Di Gu Pi	22.31	–
Yi Yi Ren	22.41	–
Lu Gen	23.01	–
Bai Zi Ren	23.10	–
Wu Wei Zi	23.28	–
Zhi Gan Cao	23.49	–
Gou Qi Zi	23.76	–
Mu Zei	23.92	–
Ma Huang Gen	24.61	–
Tai Zi Shen	25.31	–
Zhu Ru	25.42	–
Pu Gong Ying	26.16	–
Chuang Mu Xiang	26.37	–
Ze Xie	26.44	–
Ren Shen	26.82	–
Tao Ren	27.20	–
Yu Zhu	27.51	–
Fu Xiao Mai	27.83	–
Xie Bai	28.34	–
Ci Wu Jia	29.64	–
Dang Gui Wei	30.37	–
San Qi	30.89	–
Mang Xiao	31.36	–
Ze Lan	31.65	–
Hong Hua	31.94	–
Fo Shou	32.60	–
Chi Shao (Yao)	33.28	–
Ban Xia (Jiang)	33.42	–
Dang Gui	33.71	–
Sang Zhi	33.96	–
Ban Zhi Lian	35.26	–
Long Yan Rou	36.13	–
Jie Geng	36.17	–
Sang Ye	36.30	–
Cang Er Zi	37.18	–
Shan Yu Rou	37.50	–
Xi Xian Cao	37.56	–
Ku Shen	38.53	–
Huang Qin	39.18	–
Lai Fu Zi	39.62	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	40.67	–
Fu Shen	42.95	–
Jiang Huang	42.96	–
Nü Zhen Zi	44.03	–
Bai Zhu	47.40	–
(Shi) Chang Pu	48.04	–
Zhi Mu	48.43	–
Wu Mei	49.99	–
Mu Gua	50.37	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chai Hu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62661	62661	0.00	6.58
62662	62662	0.00	5.69
62905	62905	0.00	4.81
62906	62906	0.00	5.30

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Che Qian Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60255-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Che Qian Zi; Plantaginis semen

Special notes

When selecting the *Che Qian Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Che Qian Zi	2	0	4

Second-stage model

For differentiation of the substance/substance group *Che Qian Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Che Qian Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Che Qian Zi	G192HS120SH1	62607	40	from supplier
PhytoComm	Che Qian Zi	G192HS120SH1	62608	40	from supplier
PhytoComm	Che Qian Zi	G192HS120SR1	62841	40	from supplier
PhytoComm	Che Qian Zi	G192HS120SR1	62842	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Che Qian Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Che Qian Zi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Che Qian Zi	G192HS120SH1	62607†	20
PhytoComm	Che Qian Zi	G192HS120SH1	62608†	20
PhytoComm	Che Qian Zi	G192HS120SR1	62841†	20
PhytoComm	Che Qian Zi	G192HS120SR1	62842†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 6 *Apo-Ident* customers from 4 batches from the substance/substance group *Che Qian Zi*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Che Qian Zi	410704805	1
PhytoComm	Che Qian Zi	G192H0723321	2
Phytocomm	Che Qian Zi	G192H0723421	2
PhytoComm	Che Qian Zi	G192H0723621	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Che Qian Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Che Qian Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	6	851

The substance/substance group *Che Qian Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Che Qian Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hou Po	5.54	–
Yu Jin	6.87	–
Mao Dong Qing	7.45	–
Ling Zhi	7.86	–
Ren Dong Teng	8.34	–
Bai Xian Pi	9.05	–
Yin Yang Huo	9.06	–
Chai Hu	9.45	–
Dan Dou Chi	9.99	–
Du Zhong	10.39	–
Yan Hu Suo	10.86	–
Zhu Ling	11.18	–
Guang Huo Xiang	11.34	–
Shen Qu	11.52	–
Ji Li	11.71	–
Ce Bai Ye	11.84	–
Sha Ren	11.95	–
Jing Jie	12.98	–
Qiang Huo	13.30	–
Ye Jiao Teng	13.68	–
Pi Pa Ye	13.94	–
Fu Zi	13.95	–
Tu Fu Ling	14.29	–
Shan Yao	14.32	–
Fu Pen Zi	14.48	–
Chuan Xiong	14.48	–
Dan Shen	16.10	–
Zi Hua Di Ding	16.15	–
Lian Qiao	16.39	–
Ji Xue Teng	16.49	–
Bo He	16.55	–
Lian Zi	17.07	–
Bai Jiang Cao	17.64	–
She Gan	17.76	–
Gu Sui Bu	17.93	–
He Huan Pi	18.33	–
Gou Teng	19.90	–
Qing Pi	20.00	–
Qing Hao	20.17	–
Ma Huang Gen	20.45	–
Tian Hua Fen	20.50	–
Hong Jing Tian	20.65	–
Zhe Bei Mu	20.90	–
Zhi Ke	21.02	–
Jiao Gu Lan	21.11	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Ling	21.59	–
Rou Gui	21.70	–
Jin Yin Hua	21.70	–
Sheng Jiang	21.87	–
Huang Lian	22.04	–
(Fen) Bi Xie	22.14	–
Ma Huang	22.15	–
Suan Zao Ren	22.40	–
Mu Zei	22.42	–
Huo Ma Ren	22.47	–
Yi Yi Ren	22.63	–
Bai Shao Yao	22.79	–
Huang Bai	23.64	–
Gui Zhi	24.13	–
Gua Lou	24.44	–
Wu Wei Zi	24.44	–
Lu Gen	24.61	–
Gan Cao	25.46	–
Ren Shen	26.49	–
Pu Gong Ying	26.85	–
Ban Lan Gen	27.43	–
Tao Ren	28.43	–
Tai Zi Shen	28.53	–
Di Gu Pi	28.77	–
Yuan Zhi	29.15	–
Ze Lan	29.43	–
Bai Zi Ren	29.98	–
Cang Zhu	30.34	–
Mang Xiao	30.42	–
Fu Xiao Mai	31.08	–
Zhu Ru	31.70	–
Ze Xie	31.93	–
Zhi Gan Cao	32.23	–
Ban Zhi Lian	32.63	–
Chen Pi	33.28	–
Ban Xia (Jiang)	34.14	–
Ci Wu Jia	35.25	–
Xi Xian Cao	35.55	–
Gou Qi Zi	35.84	–
E Zhu	36.59	–
Lai Fu Zi	37.07	–
Fo Shou	37.32	–
Fu Shen	37.45	–
Yu Zhu	37.95	–
Chi Shao (Yao)	38.61	–
Sang Zhi	38.86	–
Chuang Mu Xiang	39.43	–
Jiang Huang	39.45	–
Jie Geng	39.81	–
San Qi	40.08	–
Huang Qin	40.55	–
Xie Bai	40.89	–
Dang Gui Wei	41.75	–
Dang Gui	42.33	–
Sang Ye	42.51	–
Nü Zhen Zi	42.82	–
Ku Shen	42.82	–
Hong Hua	43.85	–
Cang Er Zi	45.40	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Shen (Bei)	46.05	–
Shan Yu Rou	46.30	–
(Shi) Chang Pu	48.15	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Che Qian Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62607	62607	0.00	5.69
62608	62608	0.00	5.54
62841	62841	0.00	8.67
62842	62842	0.00	7.45

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chen Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60009-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chen Pi; Citri reticulatae pericarpium

Special notes

When selecting the *Chen Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chen Pi	3	0	2

Second-stage model

For differentiation of the substance/substance group *Chen Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chen Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chen Pi	G073H1128821	62377	40	from supplier
PhytoComm	Chen Pi	G073H1128821	62378	40	from supplier
PhytoComm	Chen Pi	G073HS254SG1	62669	40	from supplier
PhytoComm	Chen Pi	G073HS254SG1	62670	40	from supplier
PhytoComm	Chen Pi	G073H1128022	62909	40	from supplier
PhytoComm	Chen Pi	G073H1128022	62910	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Chen Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Chen Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chen Pi	G073H1128821	62377†	20
PhytoComm	Chen Pi	G073H1128821	62378†	20
PhytoComm	Chen Pi	G073HS254SG1	62669†	20
PhytoComm	Chen Pi	G073HS254SG1	62670†	20
PhytoComm	Chen Pi	G073H1128022	62909†	20
PhytoComm	Chen Pi	G073H1128022	62910†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Chen Pi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chen Pi	g073h1128222	1
Phytocomm	Chen Pi	G073H1128222	1
PhytoComm	Chen Pi	G073H1128222	1
Phytocomm	Chen Pi	g07h1128221	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chen Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chen Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	1	2	2	852

The substance/substance group *Chen Pi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	99.9225 % (> 99.3359 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chen Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chuan Xiong	6.50	–
Gou Qi Zi	7.82	–
(Shi) Chang Pu	7.97	–
Cang Zhu	8.57	–
Ze Xie	9.19	–
Gua Lou	9.52	–
Yuan Zhi	10.21	–
Pi Pa Ye	11.01	–
Bai Shao Yao	11.17	–
Dang Gui	11.21	–
Dang Gui Wei	11.22	–
Yu Zhu	11.28	–
Ju Hua	11.87	–
Tian Hua Fen	12.69	–
Mu Dan Pi	12.92	–
Shen Qu	13.06	–
Wu Yao	13.22	–
Shan Yu Rou	13.26	–
Suan Zao Ren	13.64	–
Dan Dou Chi	14.03	–
Shan Yao	14.04	–
Gu Sui Bu	14.18	–
Ren Dong Teng	14.26	–
Hong Jing Tian	14.72	–
Long Yan Rou	14.74	–
Mu Zei	14.80	–
Zhe Bei Mu	15.23	–
Ban Lan Gen	15.37	–
Chuang Mu Xiang	15.50	–
Lian Zi	15.55	–
Fo Shou	15.74	–
San Qi	16.22	–
Lian Qiao	16.32	–
Jie Geng	16.58	–
Xie Bai	16.74	–
Tai Zi Shen	16.79	–
Jin Yin Hua	16.82	–
Ling Zhi	16.94	–
Ji Li	17.05	–
Tu Fu Ling	17.36	–
Ye Jiao Teng	17.99	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Zi Ren	18.37	–
Fu Pen Zi	18.43	–
Dan Shen	18.53	–
Ku Shen	18.70	–
Fu Zi	18.73	–
Jiao Gu Lan	18.95	–
Sang Bai Pi	19.03	–
Fang Feng	19.37	–
Ren Shen	19.46	–
Chuan Niu Xi	19.54	–
Sang Ji Shend	19.62	–
Mu Gua	19.63	–
Sha Ren	19.78	–
Lu Gen	19.88	–
He Huan Pi	19.91	–
Di Gu Pi	20.12	–
Xiang Fu	20.28	–
Tao Ren	20.50	–
Hou Po	20.55	–
Lai Fu Zi	20.66	–
Ci Wu Jia	20.71	–
Zhi Mu	21.13	–
Chai Hu	21.38	–
Huo Ma Ren	21.67	–
Mao Dong Qing	22.02	–
Sheng Jiang	22.22	–
Zhi Gan Cao	22.33	–
Ma Huang	22.78	–
She Gan	22.80	–
Zhi Ke	22.84	–
Gou Teng	22.91	–
Fu Ling	23.96	–
Bai Xian Pi	24.17	–
Jiang Huang	24.19	–
Wu Zhu Yu	24.26	–
Gui Zhi	24.31	–
Chuan Lian Zi	24.34	–
Rou Gui	24.50	–
Wu Wei Zi	24.57	–
Gan Cao	24.85	–
Yin Chen Hao	24.86	–
Zhu Ru	25.23	–
Ji Xue Teng	25.56	–
Yin Yang Huo	25.58	–
Guang Huo Xiang	26.19	–
He Shou Wu	26.41	–
Che Qian Zi	26.41	–
Mang Xiao	26.55	–
Ce Bai Ye	26.57	–
Ban Xia (Jiang)	26.58	–
(Fen) Bi Xie	26.64	–
Ba Ji Tian	28.48	–
Yu Jin	28.58	–
Yi Mu Cao	28.71	–
Xiao Hui Xiang	28.97	–
Qiang Huo	29.11	–
Huang Qin	29.24	–
Yan Hu Suo	29.47	–
Huang Qi	30.29	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang Gen	30.47	–
Sang Zhi	30.50	–
Zi Hua Di Ding	30.60	–
Fu Xiao Mai	30.88	–
Xu Duan	31.01	–
Zhu Ling	31.13	–
Dan Zhu Ye	31.99	–
Chi Shao (Yao)	32.92	–
Hong Hua	32.99	–
Jing Jie	33.02	–
Yi Yi Ren	33.23	–
Bo He	33.36	–
Tu Si Zi	33.48	–
Qing Pi	34.30	–
Ge Gen	34.71	–
Bai Zhi	34.99	–
(Bai) Dou Kou	36.18	–
Nü Zhen Zi	37.05	–
Bing Lang	37.13	–
Huang Bai	37.27	–
Bai Hua She She Cao	37.85	–
Cang Er Zi	38.11	–
Bu Gu Zhi	38.16	–
Xin Yi	38.68	–
Huang Lian	39.07	–
E Zhu	39.08	–
(Huai) Niu Xi	39.27	–
Niu Bang Zi	39.84	–
Xuan Fu Hua	40.77	–
Hu Zhang	41.02	–
Zi Su Zi	41.66	–
Fu Shen	42.01	–
Du Zhong	43.19	–
Bai Zhu	43.44	–
Mai Men Dong	44.61	–
Sang Ye	45.23	–
Du Huo	45.50	–
Mi Huan Jun	46.07	–
Pu Gong Ying	46.15	–
Ze Lan	47.82	–
Gan Jiang	50.78	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chen Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62377	62377	0.00	12.92
62378	62378	0.00	13.78
62669	62669	0.00	6.50
62670	62670	0.00	6.58
62909	62909	0.00	8.33
62910	62910	0.00	7.97

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chi Shao (Yao)**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60014-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chi Shao (Yao); Paeoniae radix rubra

Special notes

When selecting the *Chi Shao (Yao)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chi Shao (Yao)	3	0	4

Second-stage model

For differentiation of the substance/substance group *Chi Shao (Yao)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chi Shao (Yao)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chi Shao (Yao)	G180H0705821	62363	40	from supplier
PhytoComm	Chi Shao (Yao)	G180H0705821	62364	40	from supplier
PhytoComm	Chi Shao (Yao)	G180H0705822	62481	40	from supplier
PhytoComm	Chi Shao (Yao)	G180H0705822	62482	40	from supplier
PhytoComm	Chi Shao (Yao)	G180H0705922	62831	40	from supplier
PhytoComm	Chi Shao (Yao)	G180H0705922	62832	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Chi Shao (Yao)*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Chi Shao (Yao)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chi Shao (Yao)	G180H0705821	62363†	20
PhytoComm	Chi Shao (Yao)	G180H0705821	62364†	20
PhytoComm	Chi Shao (Yao)	G180H0705822	62481†	20
PhytoComm	Chi Shao (Yao)	G180H0705822	62482†	20
PhytoComm	Chi Shao (Yao)	G180H0705922	62831†	20
PhytoComm	Chi Shao (Yao)	G180H0705922	62832†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 4 batches from the substance/substance group *Chi Shao (Yao)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chi Shao (Yao)	g180h0705122	1
Phytocomm	Chi Shao (Yao)	G180H0705321	1
Phytocomm	Chi Shao (Yao)	G180H0705422	1
Phytocomm	Chi Shao (Yao)	G189H0705321	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chi Shao (Yao)* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chi Shao (Yao)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	2	2	853

The substance/substance group *Chi Shao (Yao)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chi Shao (Yao)* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Huang Qi	10.99	–
Zhi Gan Cao	20.03	–
Mu Gua	22.48	–
Bing Lang	24.71	–
Gan Cao	25.06	–
Yuan Zhi	25.14	–
Mang Xiao	26.00	–
(Huai) Niu Xi	28.37	–
Bai Zhu	28.96	–
Lian Qiao	29.64	–
Zhi Ke	30.20	–
Xiang Fu	30.53	–
Chuan Lian Zi	30.79	–
Sang Zhi	31.99	–
Jie Geng	32.79	–
Di Gu Pi	34.79	–
Ma Huang	34.90	–
Gou Qi Zi	35.81	–
Zi Su Zi	35.83	–
Gua Lou	36.54	–
Shan Yao	37.13	–
Dang Gui	38.18	–
Ban Zhi Lian	38.23	–
Huang Qin	38.60	–
(Shi) Chang Pu	38.61	–
Chuan Niu Xi	39.21	–
Bai Zhi	40.39	–
Tian Hua Fen	40.53	–
Du Huo	40.77	–
Ju Hua	40.82	–
Fu Zi	41.93	–
Jin Yin Hua	42.08	–
Shan Yu Rou	42.14	–
Wu Wei Zi	42.23	–
Jiao Gu Lan	42.42	–
Huang Bai	43.09	–
Cang Zhu	43.14	–
Dan Shen	43.51	–
Suan Zao Ren	43.57	–
Chuan Mu Tong	43.84	–
Chai Hu	44.14	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Niu Bang Zi	44.38	–
Hong Hua	45.05	–
Zi Hua Di Ding	45.14	–
Mao Dong Qing	45.50	–
Huang Lian	45.52	–
Bo He	46.08	–
Pu Gong Ying	46.35	–
Cang Er Zi	46.36	–
Long Dan (Cao)	46.47	–
Tu Fu Ling	47.37	–
Zhe Bei Mu	47.74	–
Sha Shen (Bei)	47.75	–
Ku Shen	47.83	–
Qin Jiao	47.85	–
E Zhu	48.79	–
Lian Zi	48.86	–
Yan Hu Suo	49.20	–
Sha Ren	49.42	–
Chen Pi	49.78	–
Gou Teng	49.85	–
Yin Yang Huo	50.11	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chi Shao (Yao)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62363	62363	0.00	11.38
62364	62364	0.00	10.99
62481	62481	0.00	16.48
62482	62482	0.00	16.21
62831	62831	0.00	15.83
62832	62832	0.00	15.93

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chuan Lian Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60078-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chuan Lian Zi; Meliae toosendan fructus

Special notes

When selecting the *Chuan Lian Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chuan Lian Zi	2	0	1

Second-stage model

For differentiation of the substance/substance group *Chuan Lian Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chuan Lian Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chuan Lian Zi	G157H0306921	62893	40	from supplier
PhytoComm	Chuan Lian Zi	G157H0306921	62894	40	from supplier
PhytoComm	Chuan Lian Zi	G157H0306021	62995	40	from supplier
PhytoComm	Chuan Lian Zi	G157H0306021	62996	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Chuan Lian Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Chuan Lian Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuan Lian Zi	G157H0306921	62893†	20
PhytoComm	Chuan Lian Zi	G157H0306921	62894†	20
PhytoComm	Chuan Lian Zi	G157H0306021	62995†	20
PhytoComm	Chuan Lian Zi	G157H0306021	62996†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Chuan Lian Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chuan Lian Zi	G157H0306421	1
PhytoComm	Chuan Lian Zi	G157H0306421	2

- 854 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chuan Lian Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chuan Lian Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	4	0	3	850

The substance/substance group *Chuan Lian Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.4961 % (> 98.9101 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chuan Lian Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sang Zhi	6.88	–
Tian Hua Fen	8.63	–
Di Gu Pi	9.99	–
Jie Geng	11.03	–
Sha Shen (Bei)	11.45	–
Bai He	11.64	–
Gan Jiang	13.01	–
Mai Ya	13.25	–
Long Dan (Cao)	14.32	–
Chuan Mu Tong	15.36	–
Bai Zhu	16.28	–
E Zhu	17.02	–
Lai Fu Zi	17.22	–
Zi Su Zi	17.29	–
Mi Huan Jun	17.42	–
Qin Jiao	18.43	–
Dang Gui	19.97	–
Ji Li	20.11	–
Zhi Gan Cao	20.18	–
Huang Qi	20.26	–
(Huai) Niu Xi	20.93	–
Cang Er Zi	22.17	–
(Shi) Chang Pu	22.24	–
Da Zao	22.66	–
Mang Xiao	23.14	–
Gua Lou	24.33	–
Yan Hu Suo	24.83	–
Shan Yao	26.21	–
Bai Zhi	28.95	–
Jiang Huang	29.06	–
Niu Bang Zi	29.22	–
Lian Qiao	29.67	–
Ban Zhi Lian	31.59	–
Chuan Niu Xi	31.85	–
Chi Shao (Yao)	32.22	–
Mu Gua	34.32	–
Yuan Zhi	35.05	–
Lian Zi	35.64	–
Ku Shen	35.75	–
Xiao Hui Xiang	36.53	–
Suan Zao Ren	37.33	–
(Bai) Dou Kou	37.43	–
Chen Pi	37.54	–
Fu Ling	37.82	–
Sha Ren	38.18	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Xian Pi	38.50	–
Wu Wei Zi	39.74	–
Tu Fu Ling	40.07	–
Bai Shao Yao	40.28	–
Chai Hu	40.42	–
Jin Yin Hua	41.61	–
He Huan Pi	41.86	–
Pu Gong Ying	41.94	–
Qiang Huo	42.11	–
Bo He	43.30	–
Bing Lang	43.46	–
Ren Dong Teng	43.73	–
Rou Gui	44.16	–
Zhi Ke	44.74	–
Zi Hua Di Ding	45.31	–
Chuan Xiong	45.46	–
Ye Jiao Teng	46.22	–
Gou Teng	46.41	–
Yi Yi Ren	46.55	–
Ce Bai Ye	46.60	–
Shen Qu	46.95	–
Du Huo	47.05	–
Gan Cao	47.21	–
Huang Lian	47.26	–
Xiang Fu	47.42	–
Dan Dou Chi	47.54	–
Yu Jin	48.27	–
Fang Feng	48.41	–
Fo Shou	48.69	–
Dong Gua Zi	48.96	–
Zhu Ling	49.18	–
Shan Yu Rou	49.74	–
Mao Dong Qing	50.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chuan Lian Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62893	62893	0.00	8.63
62894	62894	0.00	8.16
62995	62995	0.00	7.01
62996	62996	0.00	6.88

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chuan Mu Tong**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60173-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chuan Mu Tong; Clematidis armandii caulis

Special notes

When selecting the *Chuan Mu Tong* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chuan Mu Tong	1	0	2

Second-stage model

For differentiation of the substance/substance group *Chuan Mu Tong* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chuan Mu Tong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chuan Mu Tong	G009H0310921	62947	40	from supplier
PhytoComm	Chuan Mu Tong	G009H0310921	62948	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Chuan Mu Tong*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Chuan Mu Tong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuan Mu Tong	G009H0310921	62947†	20
PhytoComm	Chuan Mu Tong	G009H0310921	62948†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Chuan Mu Tong*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chuan Mu Tong	g009h0310321	1
Phytocomm	Chuan Mu Tong	G009H0310321	1
PhytoComm	Chuan Mu Tong	G009H0310521	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chuan Mu Tong* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chuan Mu Tong* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	80	0	24 519
Type B	0	40	0	12 437
Type C	0	0	3	854

The substance/substance group *Chuan Mu Tong* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9779 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chuan Mu Tong* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mi Huan Jun	4.89	–
Sang Zhi	11.09	–
Dang Gui	11.85	–
Di Gu Pi	12.35	–
Chuan Lian Zi	13.39	–
Bai He	13.57	–
Mai Ya	14.21	–
Sha Shen (Bei)	18.10	–
Gan Jiang	18.25	–
E Zhu	19.17	–
Qin Jiao	19.20	–
Tian Hua Fen	19.33	–
Long Dan (Cao)	19.47	–
Ji Li	21.24	–
Ban Zhi Lian	22.35	–
(Huai) Niu Xi	24.19	–
Mang Xiao	24.94	–
Lai Fu Zi	25.56	–
Huang Qi	27.57	–
Yan Hu Suo	27.73	–
Chuan Niu Xi	27.77	–
(Shi) Chang Pu	29.02	–
Da Zao	29.38	–
Bai Zhi	29.39	–
Zhi Gan Cao	29.62	–
Zi Su Zi	29.66	–
Jie Geng	29.86	–
Jiang Huang	30.31	–
Bai Zhu	31.12	–
Shan Yao	31.70	–
Cang Er Zi	32.27	–
(Bai) Dou Kou	34.99	–
Sha Ren	36.63	–
Bing Lang	37.80	–
Niu Bang Zi	38.10	–
Chen Pi	38.13	–
Lian Qiao	38.15	–
He Huan Pi	38.88	–
Fu Ling	39.66	–
Rou Gui	39.94	–
Lian Zi	40.13	–
Gua Lou	40.36	–
Bai Xian Pi	41.90	–
Chi Shao (Yao)	43.56	–
Bai Shao Yao	43.73	–
Mu Dan Pi	44.16	–
Yi Yi Ren	44.49	–
Xiao Hui Xiang	45.40	–
Wu Wei Zi	45.80	–
Suan Zao Ren	46.01	–
Xiang Fu	46.10	–
Tu Fu Ling	46.37	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mu Gua	46.63	–
Ku Shen	46.74	–
Chai Hu	47.69	–
Zhi Ke	48.47	–
Ji Xue Teng	48.55	–
Yuan Zhi	48.88	–
Fang Feng	48.91	–
Ye Jiao Teng	49.07	–
Shan Yu Rou	49.18	–
Jin Yin Hua	49.35	–
Zi Hua Di Ding	49.49	–
Huang Lian	50.23	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chuan Mu Tong* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62947	62947	0.00	4.89
62948	62948	0.00	7.38

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chuan Niu Xi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60122-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chuan Niu Xi; Cyathulae radix

Special notes

When selecting the *Chuan Niu Xi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chuan Niu Xi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Chuan Niu Xi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chuan Niu Xi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chuan Niu Xi	G318H0431923	62817	40	from supplier
PhytoComm	Chuan Niu Xi	G318H0431923	62818	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Chuan Niu Xi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Chuan Niu Xi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuan Niu Xi	G318H0431923	62817†	20
PhytoComm	Chuan Niu Xi	G318H0431923	62818†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Chuan Niu Xi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chuan Niu Xi	g318h0431321	1
PhytoComm	Chuan Niu Xi	G318H0431321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chuan Niu Xi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chuan Niu Xi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	4	2	0	851

The substance/substance group *Chuan Niu Xi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.5194 % (> 98.9345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chuan Niu Xi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui	12.81	–
Du Huo	13.76	–
Sang Zhi	14.53	–
Chuan Lian Zi	14.78	–
(Huai) Niu Xi	17.17	–
Shan Yu Rou	18.21	–
Ba Ji Tian	18.93	–
Ji Li	19.30	–
Chen Pi	19.76	–
Xiang Fu	22.42	–
(Shi) Chang Pu	22.54	–
Huang Qi	23.48	–
Fang Feng	23.81	–
Zhi Gan Cao	24.52	–
E Zhu	24.55	–
Long Dan (Cao)	24.71	–
Da Zao	25.38	–
Bing Lang	25.46	–
Mang Xiao	26.53	–
Lai Fu Zi	27.15	–
Ju Hua	27.98	–
Bai Zhu	28.40	–
Zi Su Zi	29.58	–
Jie Geng	29.58	–
Tian Hua Fen	29.65	–
Mu Dan Pi	29.67	–
He Huan Pi	29.88	–
Yan Hu Suo	30.18	–
Mi Huan Jun	30.53	–
Jiang Huang	31.14	–
Yuan Zhi	33.18	–
Qin Jiao	34.30	–
Shan Yao	34.43	–
Lian Qiao	34.55	–
Sha Shen (Bei)	35.70	–
Di Gu Pi	35.83	–
Wu Yao	36.23	–
Gua Lou	36.68	–
Sha Ren	37.14	–
Niu Bang Zi	37.82	–
Mai Ya	38.73	–
Chuan Mu Tong	40.10	–
Bai Zhi	40.15	–
Bai He	40.95	–
(Bai) Dou Kou	41.15	–
Xiao Hui Xiang	41.33	–
Gan Jiang	41.72	–
Ku Shen	42.01	–
Bai Hua She She Cao	42.56	–
Gu Sui Bu	42.91	–
Chi Shao (Yao)	43.06	–
Zhi Ke	43.48	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dong Gua Zi	44.59	–
Xu Duan	44.75	–
Cang Er Zi	45.78	–
Suan Zao Ren	47.12	–
Sang Ji Shend	47.18	–
Lian Zi	47.78	–
Jin Yin Hua	48.45	–
Wu Wei Zi	49.51	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chuan Niu Xi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62817	62817	0.00	12.81
62818	62818	0.00	12.92

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chuan Xiong**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60013-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chuan Xiong; Ligustici wallichii radix

Special notes

When selecting the *Chuan Xiong* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chuan Xiong	1	0	3

Second-stage model

For differentiation of the substance/substance group *Chuan Xiong* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chuan Xiong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chuan Xiong	G136HS031SH1	62689	40	from supplier
PhytoComm	Chuan Xiong	G136HS031SH1	62690	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Chuan Xiong*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Chuan Xiong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuan Xiong	G136HS031SH1	62689†	20
PhytoComm	Chuan Xiong	G136HS031SH1	62690†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 6 *Apo-Ident* customers from 3 batches from the substance/substance group *Chuan Xiong*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chuan Xiong	g136h0305121	1
Phytocomm	Chuan Xiong	G136H0305221	1
PhytoComm	Chuan Xiong	G136H0305221	2
Phytocomm	Chuan Xiong	G136H0305421	2

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chuan Xiong* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chuan Xiong* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	39	1	12 437
Type C	0	0	6	851

The substance/substance group *Chuan Xiong* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chuan Xiong* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Pi Pa Ye	4.86	–
Shan Yao	5.63	–
Shen Qu	6.76	–
Dan Dou Chi	7.29	–
Chen Pi	9.06	–
Cang Zhu	9.30	–
Ren Dong Teng	9.38	–
Tian Hua Fen	9.75	–
Gu Sui Bu	9.92	–
Hong Jing Tian	9.94	–
Jin Yin Hua	10.16	–
Dan Shen	10.84	–
Ban Lan Gen	10.86	–
Ji Li	11.30	–
Lian Qiao	11.35	–
Gua Lou	11.45	–
Guang Huo Xiang	11.47	–
Hou Po	11.48	–
Ze Xie	11.81	–
Lian Zi	11.98	–
Bai Shao Yao	12.92	–
Fu Pen Zi	12.97	–
Fu Zi	13.02	–
Gou Qi Zi	13.08	–
Ling Zhi	13.15	–
Yuan Zhi	13.66	–
Suan Zao Ren	13.70	–
Mu Zei	13.95	–
Zhe Bei Mu	14.19	–
He Huan Pi	14.32	–
Zhi Gan Cao	14.38	–
Mao Dong Qing	14.43	–
Jiao Gu Lan	14.95	–
Chai Hu	15.08	–
She Gan	15.27	–
Yu Zhu	15.88	–
Bai Zi Ren	16.88	–
Tai Zi Shen	17.15	–
Fu Ling	17.65	–
Tu Fu Ling	17.85	–
Lu Gen	17.93	–
Ce Bai Ye	18.01	–
Gan Cao	18.12	–
Che Qian Zi	18.40	–
Dang Gui	18.59	–
Bai Xian Pi	18.73	–
Huo Ma Ren	18.82	–
Dang Gui Wei	18.85	–
Ye Jiao Teng	18.98	–
Yin Yang Huo	19.08	–
Sheng Jiang	19.20	–
Gou Teng	19.61	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Jin	19.91	–
Chuang Mu Xiang	20.17	–
Ku Shen	20.39	–
Qiang Huo	20.44	–
Gui Zhi	20.63	–
Di Gu Pi	20.76	–
Xie Bai	20.80	–
Ma Huang	20.85	–
Ci Wu Jia	20.99	–
San Qi	21.13	–
Yan Hu Suo	21.33	–
Long Yan Rou	21.88	–
Zhu Ling	21.89	–
Ji Xue Teng	21.92	–
(Fen) Bi Xie	22.14	–
Tao Ren	22.21	–
Ren Shen	22.29	–
Zi Hua Di Ding	22.57	–
Zhi Ke	22.77	–
Wu Wei Zi	23.00	–
Jie Geng	23.11	–
Rou Gui	24.10	–
Fo Shou	24.44	–
Ma Huang Gen	24.85	–
Zhu Ru	25.32	–
Qing Pi	25.32	–
Lai Fu Zi	25.80	–
Bo He	26.56	–
Jing Jie	27.06	–
Zhi Mu	27.99	–
Shan Yu Rou	28.35	–
Huang Bai	28.68	–
Fu Xiao Mai	29.03	–
Ban Xia (Jiang)	29.55	–
Sha Ren	29.80	–
Huang Qin	30.15	–
Yi Yi Ren	30.32	–
Mu Gua	31.37	–
Mang Xiao	31.98	–
Sang Zhi	32.14	–
Hong Hua	32.75	–
Huang Lian	32.92	–
Du Zhong	33.44	–
Chi Shao (Yao)	35.43	–
Cang Er Zi	36.59	–
Fu Shen	41.50	–
Bai Jiang Cao	42.01	–
Qing Hao	44.90	–
Sang Ye	49.65	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chuan Xiong* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62689	62689	0.00	5.37
62690	62690	0.00	4.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Chuang Mu Xiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50323-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Chuang Mu Xiang; Vladimirae radix

Special notes

When selecting the *Chuang Mu Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Chuang Mu Xiang	2	0	3

Second-stage model

For differentiation of the substance/substance group *Chuang Mu Xiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Chuang Mu Xiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Chuang Mu Xiang	G219HS046SK1	62643	40	from supplier
PhytoComm	Chuang Mu Xiang	G219HS046SK1	62644	40	from supplier
PhytoComm	Chuang Mu Xiang	G219HS046SR1	62861	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Chuang Mu Xiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 480 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 119 spectra of 4 reference samples from the substance/substance group *Chuang Mu Xiang*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuang Mu Xiang	G219HS046SK1	62643†	20
PhytoComm	Chuang Mu Xiang	G219HS046SK1	62644†	20
PhytoComm	Chuang Mu Xiang	G219HS046SR1	62861†	20
PhytoComm	Chuang Mu Xiang	G219HS046SR1	62862	59

- 12 358 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the

substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Chuang Mu Xiang*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Chuang Mu Xiang	g219h0309321	1
Phytocomm	Chuang Mu Xiang	g219h0309322	1
Phytocomm	Chuang Mu Xiang	G219H0309322	1
Phytocomm	Chuang Mu Xiang	G219H0309521	2

- 852 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Chuang Mu Xiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Chuang Mu Xiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	120	0	24 480
Type B	0	119	0	12 358
Type C	0	0	5	852

The substance/substance group *Chuang Mu Xiang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 95.0000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 94.9580 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Chuang Mu Xiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Gou Qi Zi	9.16	–
Chen Pi	10.28	–
Xie Bai	11.17	–
Hong Jing Tian	12.36	–
Chuan Xiong	12.41	–
Cang Zhu	13.92	–
Gua Lou	14.10	–
Jin Yin Hua	14.12	–
Shan Yao	14.95	–
Yuan Zhi	14.95	–
Yu Zhu	15.92	–
Jiao Gu Lan	16.17	–
Long Yan Rou	16.71	–
Bai Shao Yao	17.73	–
Shan Yu Rou	17.75	–
Pi Pa Ye	17.84	–
Dan Shen	17.93	–
Ren Shen	17.94	–
Dang Gui	18.71	–
Dan Dou Chi	18.72	–
Suan Zao Ren	19.11	–
Tian Hua Fen	19.17	–
Ma Huang	19.42	–
Zhi Ke	19.67	–
Dang Gui Wei	19.81	–
Lian Zi	20.06	–
Gu Sui Bu	20.95	–
Lian Qiao	21.01	–
Zi Hua Di Ding	21.04	–
Mu Gua	21.32	–
Ren Dong Teng	21.59	–
Di Gu Pi	21.64	–
Hong Hua	21.92	–
Jie Geng	22.00	–
Fu Pen Zi	22.43	–
Gan Cao	22.48	–
Zhi Gan Cao	23.09	–
Shen Qu	23.11	–
Zhe Bei Mu	23.15	–
Ji Li	23.30	–
San Qi	23.34	–
Hou Po	23.39	–
Ku Shen	23.48	–
Zhi Mu	23.77	–
Ze Xie	23.78	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chai Hu	24.08	–
She Gan	24.33	–
Mu Zei	24.47	–
Tai Zi Shen	24.84	–
Lu Gen	24.89	–
Ling Zhi	25.08	–
Guang Huo Xiang	25.13	–
Mao Dong Qing	25.61	–
Wu Wei Zi	26.04	–
Ye Jiao Teng	26.23	–
Ban Lan Gen	26.45	–
Tao Ren	27.19	–
Yin Yang Huo	27.27	–
Fu Zi	27.51	–
Bai Xian Pi	27.58	–
Bai Zi Ren	27.61	–
Ji Xue Teng	27.98	–
He Huan Pi	28.00	–
Gou Teng	28.92	–
Tu Fu Ling	29.20	–
Mang Xiao	30.40	–
Huang Qin	30.98	–
Yan Hu Suo	31.02	–
Gui Zhi	31.12	–
Rou Gui	31.17	–
Sheng Jiang	31.37	–
Ci Wu Jia	31.85	–
Bo He	31.88	–
Qiang Huo	32.31	–
Fo Shou	32.37	–
Sang Zhi	32.41	–
Cang Er Zi	32.49	–
Che Qian Zi	32.94	–
Huang Bai	33.02	–
Lai Fu Zi	33.42	–
Chi Shao (Yao)	33.59	–
Huo Ma Ren	33.93	–
Ce Bai Ye	33.95	–
Sha Ren	34.04	–
Yu Jin	34.26	–
Jing Jie	34.42	–
Zhu Ru	34.55	–
Fu Ling	35.14	–
Qing Pi	35.37	–
(Fen) Bi Xie	35.59	–
Mai Men Dong	36.20	–
Ma Huang Gen	37.53	–
Chuan Lian Zi	39.70	–
Zhu Ling	40.10	–
Huang Lian	40.14	–
Du Zhong	41.95	–
Ban Xia (Jiang)	42.54	–
Huang Qi	42.54	–
Fu Xiao Mai	42.89	–
Niu Bang Zi	43.10	–
Bai Jiang Cao	44.80	–
Yi Yi Ren	46.19	–
Sang Ye	50.08	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Chuang Mu Xiang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62643	62643	0.00	10.28
62644	62644	0.00	10.39
62861	62861	0.00	9.16

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ci Wu Jia**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60354-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ci Wu Jia; Eleutherococci radix

Special notes

When selecting the *Ci Wu Jia* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ci Wu Jia	1	0	0

Second-stage model

For differentiation of the substance/substance group *Ci Wu Jia* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ci Wu Jia*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ci Wu Jia	G302HS422SH1	62623	40	from supplier
PhytoComm	Ci Wu Jia	G302HS422SH1	62624	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ci Wu Jia*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ci Wu Jia*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ci Wu Jia	G302HS422SH1	62623†	20
PhytoComm	Ci Wu Jia	G302HS422SH1	62624†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ci Wu Jia*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ci Wu Jia* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ci Wu Jia* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	72	8	24 517
Type B	1	29	11	12 436
Type C	0	0	0	857

The substance/substance group *Ci Wu Jia* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9893 % (> 99.9743 %)	90.0000 % (> 86.2500 %)
Type B	99.9964 % (> 99.9667 %)	72.5000 % (> 65.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ci Wu Jia* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Sheng Jiang	2.74	–
He Huan Pi	3.49	–
Fu Xiao Mai	4.23	–
Tao Ren	4.59	–
Gui Zhi	4.94	–
Lai Fu Zi	5.78	–
Di Gu Pi	7.53	–
Fu Ling	8.26	–
Zhu Ru	8.26	–
Ban Xia (Jiang)	8.80	–
Bai Zi Ren	8.93	–
Yi Yi Ren	9.27	–
Lian Zi	9.64	–
Rou Gui	10.06	–
Ji Li	10.51	–
Gou Teng	10.55	–
Mu Zei	10.63	–
Huo Ma Ren	10.88	–
Ling Zhi	11.74	–
Bai Shao Yao	12.09	–
Tai Zi Shen	13.07	–
Shen Qu	13.27	–
Zhe Bei Mu	13.61	–
Bai Xian Pi	13.69	–
Gu Sui Bu	15.35	–
Ji Xue Teng	16.04	–
Fo Shou	16.41	–
Ye Jiao Teng	16.51	–
Tian Hua Fen	16.81	–
Lu Gen	17.02	–
Ma Huang Gen	17.12	–
Gua Lou	17.85	–
Ze Xie	19.95	–
Ren Dong Teng	20.19	–
Suan Zao Ren	20.88	–
Yu Jin	21.27	–
Tu Fu Ling	21.33	–
Zhu Ling	21.46	–
Yuan Zhi	21.82	–
Fu Shen	22.27	–
Chuan Xiong	22.87	–
Shan Yao	22.98	–
(Fen) Bi Xie	23.16	–
Pi Pa Ye	23.93	–
Lian Qiao	24.93	–
Chen Pi	25.36	–
She Gan	25.39	–
Ce Bai Ye	26.26	–
Dan Dou Chi	28.05	–
Yan Hu Suo	28.52	–
Jin Yin Hua	29.04	–
Dan Shen	30.13	–
Mao Dong Qing	30.63	–
Cang Zhu	31.15	–
Jiao Gu Lan	31.55	–
Chai Hu	31.80	–
Fu Zi	32.03	–
Fu Pen Zi	32.20	–
Jie Geng	32.27	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dang Gui Wei	32.46	–
Zhi Mu	33.54	–
San Qi	33.73	–
Mang Xiao	34.03	–
Ren Shen	34.24	–
Yu Zhu	34.38	–
Yin Yang Huo	34.62	–
Dang Gui	35.04	–
Shan Yu Rou	35.31	–
Zhi Ke	35.41	–
Ban Lan Gen	35.54	–
Hong Jing Tian	37.38	–
Guang Huo Xiang	37.88	–
Hou Po	38.00	–
Che Qian Zi	39.51	–
Huang Qin	40.37	–
Gan Cao	41.11	–
Ma Huang	41.29	–
Mu Gua	41.69	–
Qiang Huo	41.96	–
Wu Wei Zi	42.16	–
Gou Qi Zi	45.56	–
Huang Lian	47.74	–
Chuang Mu Xiang	48.16	–
Jing Jie	48.40	–
Sang Zhi	48.72	–
Long Yan Rou	49.36	–
Zi Hua Di Ding	49.45	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ci Wu Jia* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62623	62623	0.00	2.74
62624	62624	0.00	2.85

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Da Huang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60140-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Da Huang; Rhei radix et rhizoma

Special notes

When selecting the *Da Huang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Da Huang	1	0	3

Second-stage model

For differentiation of the substance/substance group *Da Huang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Da Huang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Da Huang	G212H0320822	62495	40	from supplier
PhytoComm	Da Huang	G212H0320822	62496	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Da Huang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Da Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Da Huang	G212H0320822	62495†	20
PhytoComm	Da Huang	G212H0320822	62496†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 7 spectra from 5 *Apo-Ident* customers from 3 batches from the substance/substance group *Da Huang*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Da Huang	g212h0320023	1
Phytocomm	Da Huang	G212H0320321	2
PhytoComm	Da Huang	G212H0320321	1
Phytocomm	Da Huang	G212H0320521	3

- 850 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Da Huang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Da Huang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	1	6	850

The substance/substance group *Da Huang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8255 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Da Huang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hua Shi	36.18	–
Mang Xiao	37.52	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Da Huang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62495	62495	0.00	36.21
62496	62496	0.00	36.18

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Da Zao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60055-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Da Zao; Zizyphi jujubae fructus

Special notes

When selecting the *Da Zao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Da Zao	2	0	2

Second-stage model

For differentiation of the substance/substance group *Da Zao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Da Zao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Da Zao	G129H0316821	62539	40	from supplier
PhytoComm	Da Zao	G129H0316821	62540	40	from supplier
PhytoComm	Da Zao	G129H0316921	63007	40	from supplier
PhytoComm	Da Zao	G129H0316921	63008	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Da Zao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Da Zao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Da Zao	G129H0316821	62539†	20
PhytoComm	Da Zao	G129H0316821	62540†	20
PhytoComm	Da Zao	G129H0316921	63007†	20
PhytoComm	Da Zao	G129H0316921	63008†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Da Zao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Da Zao	g129h0316221	1
Phytocomm	Da Zao	G129H0316521	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Da Zao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Da Zao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	1	3	853

The substance/substance group *Da Zao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Da Zao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui	15.50	–
Long Dan (Cao)	20.96	–
Bing Lang	22.96	–
Mang Xiao	24.76	–
Chuan Niu Xi	26.68	–
Qin Jiao	27.04	–
Sang Zhi	27.42	–
Huang Qi	28.13	–
(Shi) Chang Pu	28.26	–
Shan Yu Rou	29.29	–
Chen Pi	29.49	–
Chuan Mu Tong	30.01	–
Mai Ya	31.13	–
He Huan Pi	32.73	–
Zhi Gan Cao	32.92	–
(Huai) Niu Xi	33.15	–
Mi Huan Jun	34.18	–
Mu Dan Pi	34.38	–
Jie Geng	34.67	–
E Zhu	35.55	–
Ba Ji Tian	36.06	–
Sha Shen (Bei)	37.04	–
Bai He	37.26	–
Tian Hua Fen	39.69	–
Fang Feng	39.89	–
Lai Fu Zi	39.90	–
Ji Li	40.11	–
Zi Su Zi	40.54	–
Chuan Lian Zi	40.60	–
Ju Hua	40.79	–
Xiang Fu	41.85	–
Yuan Zhi	41.89	–
Bai Zhi	43.25	–
Bai Zhu	44.09	–
Jiang Huang	48.53	–
Shan Yao	49.01	–
Du Huo	49.17	–
Sha Ren	50.02	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Da Zao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62539	62539	0.00	15.75
62540	62540	0.00	15.50
63007	63007	0.00	19.11
63008	63008	0.00	19.11

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dan Dou Chi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60794-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dan Dou Chi; Sojae semen praeparatum

Special notes

When selecting the *Dan Dou Chi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dan Dou Chi	2	0	1

Second-stage model

For differentiation of the substance/substance group *Dan Dou Chi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dan Dou Chi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dan Dou Chi	G232HS239SK1	62713	40	from supplier
PhytoComm	Dan Dou Chi	G232HS239SK1	62714	40	from supplier
PhytoComm	Dan Dou Chi	G232HS239TK1	62793	40	from supplier
PhytoComm	Dan Dou Chi	G232HS239TK1	62794	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Dan Dou Chi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Dan Dou Chi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dan Dou Chi	G232HS239SK1	62713†	20
PhytoComm	Dan Dou Chi	G232HS239SK1	62714†	20
PhytoComm	Dan Dou Chi	G232HS239TK1	62793†	20
PhytoComm	Dan Dou Chi	G232HS239TK1	62794†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Dan Dou Chi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Dan Dou Chi	G232HS239	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dan Dou Chi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dan Dou Chi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	6	151	9	24 434
Type B	4	67	13	12 393
Type C	0	0	1	856

The substance/substance group *Dan Dou Chi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9750 % (> 99.9600 %)	94.3750 % (> 92.5000 %)
Type B	99.9500 % (> 99.9202 %)	83.7500 % (> 80.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dan Dou Chi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Xian Pi	3.23	–
Ce Bai Ye	4.17	–
Guang Huo Xiang	4.72	–
Ren Dong Teng	6.06	–
Ling Zhi	6.08	–
Lian Zi	8.13	–
Ye Jiao Teng	8.44	–
Shen Qu	8.61	–
She Gan	9.42	–
Mao Dong Qing	9.45	–
Tu Fu Ling	9.46	–
Yin Yang Huo	10.05	–
Zhu Ling	10.25	–
(Fen) Bi Xie	10.44	–
Jing Jie	10.56	–
Hou Po	10.72	–
Gu Sui Bu	10.77	–
Yu Jin	11.05	–
Suan Zao Ren	11.10	–
Chai Hu	11.28	–
Fu Pen Zi	12.17	–
Dan Shen	12.25	–
Chuan Xiong	12.67	–
Ji Li	13.41	–
Fu Zi	13.52	–
He Huan Pi	13.61	–
Shan Yao	13.68	–
Che Qian Zi	13.69	–
Rou Gui	13.80	–
Gan Cao	14.09	–
Pi Pa Ye	14.22	–
Lu Gen	14.39	–
Qiang Huo	14.76	–
Huo Ma Ren	15.63	–
Gou Teng	15.64	–
Zi Hua Di Ding	15.71	–
Yan Hu Suo	15.82	–
Bai Shao Yao	16.18	–
Sheng Jiang	16.31	–
Ji Xue Teng	16.42	–
Fu Ling	16.46	–
Lian Qiao	16.47	–
Jin Yin Hua	17.31	–
Gua Lou	17.50	–
Tian Hua Fen	17.55	–
Gui Zhi	17.75	–
Ma Huang Gen	17.79	–
Mu Zei	18.36	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Lan Gen	19.22	–
Ma Huang	19.24	–
Hong Jing Tian	19.82	–
Zhe Bei Mu	20.18	–
Jiao Gu Lan	20.41	–
Sha Ren	20.57	–
Tao Ren	21.06	–
Yi Yi Ren	21.63	–
Bo He	22.07	–
Cang Zhu	22.12	–
Zhi Gan Cao	22.18	–
Huang Bai	22.19	–
Di Gu Pi	22.36	–
Chen Pi	23.51	–
Du Zhong	24.01	–
Bai Zi Ren	24.13	–
Tai Zi Shen	24.32	–
Zhi Ke	24.58	–
Yuan Zhi	24.72	–
Chuang Mu Xiang	24.87	–
Ci Wu Jia	25.33	–
Ze Xie	25.97	–
Qing Pi	26.38	–
Yu Zhu	26.53	–
Fo Shou	26.81	–
Fu Xiao Mai	26.97	–
Gou Qi Zi	27.59	–
Ban Xia (Jiang)	28.75	–
Huang Lian	29.04	–
Wu Wei Zi	29.41	–
Zhu Ru	29.98	–
Ren Shen	30.02	–
Dang Gui Wei	30.33	–
Mang Xiao	32.29	–
Qing Hao	32.59	–
Dang Gui	32.61	–
Fu Shen	34.57	–
Bai Jiang Cao	34.63	–
San Qi	35.10	–
Jie Geng	35.41	–
Sang Zhi	35.69	–
Lai Fu Zi	36.51	–
Xie Bai	36.90	–
Huang Qin	37.10	–
Long Yan Rou	37.29	–
Chi Shao (Yao)	39.12	–
Ban Zhi Lian	39.37	–
Hong Hua	41.65	–
Shan Yu Rou	41.78	–
Ku Shen	44.33	–
Zhi Mu	45.81	–
Cang Er Zi	45.89	–
Bai Zhu	47.37	–
Pu Gong Ying	47.79	–
Mu Gua	48.98	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dan Dou Chi* is separated from critical neighbours in a second-stage model,

all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62713	62713	0.00	3.23
62714	62714	0.00	4.12
62793	62793	0.00	4.17
62794	62794	0.00	4.39

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dan Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60030-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dan Shen; Salviae miltiorrhizae radix

Special notes

When selecting the *Dan Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dan Shen	1	0	4

Second-stage model

For differentiation of the substance/substance group *Dan Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dan Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dan Shen	G214HS059SG1	62573	40	from supplier
PhytoComm	Dan Shen	G214HS059SG1	62574	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Dan Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Dan Shen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dan Shen	G214HS059SG1	62573†	20
PhytoComm	Dan Shen	G214HS059SG1	62574†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 9 spectra from 5 *Apo-Ident* customers from 5 batches from the substance/substance group *Dan Shen*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Dan Shen	g214h0441121	1
Phytocomm	Dan Shen	G214H0441121	1
Phytocomm	Dan Shen	G214H0441221	1
PhytoComm	Dan Shen	G214H0441221	2
Phytocomm	Dan Shen	G214H0441322	1
PhytoComm	Dan Shen	G214H0441322	1
Phytocomm	Dan Shen	G214H0441522	1
PhytoComm	Dan Shen	G214H0441522	1

- 848 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dan Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dan Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	9	848

The substance/substance group *Dan Shen* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8251 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dan Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	4.63	–
Fu Zi	8.55	–
Huang Bai	9.01	–
Bo He	9.14	–
Gan Cao	10.27	–
Hou Po	10.64	–
Jin Yin Hua	10.64	–
Hong Jing Tian	10.92	–
Yin Yang Huo	11.52	–
Chai Hu	11.58	–
Qiang Huo	11.60	–
Zhi Ke	11.69	–
Fu Pen Zi	11.80	–
Jing Jie	12.36	–
Ye Jiao Teng	12.41	–
Jiao Gu Lan	12.82	–
Sha Ren	13.05	–
Dan Dou Chi	14.04	–
Ce Bai Ye	15.28	–
Ma Huang	15.49	–
Chuan Xiong	16.31	–
Yu Jin	16.43	–
Shan Yao	17.19	–
Mao Dong Qing	17.63	–
Bai Xian Pi	17.63	–
Zhi Gan Cao	17.87	–
Ren Dong Teng	18.11	–
Du Zhong	18.29	–
Pi Pa Ye	18.35	–
Guang Huo Xiang	18.57	–
Ji Xue Teng	18.64	–
Huang Lian	18.91	–
Suan Zao Ren	18.91	–
Tu Fu Ling	19.15	–
Qing Pi	19.24	–
Chuang Mu Xiang	19.39	–
Tian Hua Fen	19.57	–
Wu Wei Zi	20.24	–
Lian Zi	20.27	–
Bai Shao Yao	20.34	–
Bai Jiang Cao	20.42	–
Gu Sui Bu	21.45	–
Yan Hu Suo	21.47	–
Che Qian Zi	21.58	–
She Gan	21.64	–
Di Gu Pi	22.11	–
Gua Lou	22.53	–
Hong Hua	22.55	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qing Hao	22.64	–
Lu Gen	22.67	–
Ban Lan Gen	22.83	–
Gou Qi Zi	23.15	–
Shen Qu	23.45	–
(Fen) Bi Xie	24.59	–
Lian Qiao	24.62	–
Cang Zhu	24.88	–
He Huan Pi	25.16	–
Ji Li	25.30	–
Yuan Zhi	26.12	–
Ling Zhi	26.20	–
Mu Zei	26.96	–
Ren Shen	27.19	–
Chi Shao (Yao)	27.31	–
Gou Teng	28.27	–
Zhu Ling	28.90	–
Pu Gong Ying	29.47	–
Mang Xiao	29.93	–
Rou Gui	30.19	–
Tao Ren	31.44	–
Cang Er Zi	31.82	–
Huang Qin	31.82	–
Sang Ye	32.00	–
Zhe Bei Mu	32.02	–
Huo Ma Ren	32.67	–
Jie Geng	33.04	–
Ban Zhi Lian	33.10	–
Chen Pi	33.15	–
Sheng Jiang	33.20	–
Ze Lan	33.57	–
Shan Yu Rou	34.58	–
Ma Huang Gen	35.03	–
Fu Ling	36.29	–
Xie Bai	36.86	–
Dang Gui	36.94	–
Sang Zhi	37.64	–
Gui Zhi	38.05	–
Ze Xie	38.05	–
Tai Zi Shen	38.11	–
Ku Shen	38.85	–
Dang Gui Wei	39.21	–
Long Yan Rou	39.31	–
Yu Zhu	39.41	–
San Qi	41.15	–
Ci Wu Jia	41.57	–
Bai Zi Ren	43.15	–
Ban Xia (Jiang)	43.34	–
Xi Xian Cao	44.15	–
Yi Yi Ren	44.41	–
Fo Shou	45.57	–
(Shi) Chang Pu	46.55	–
E Zhu	46.61	–
Niu Bang Zi	47.10	–
Zi Su Zi	47.28	–
Jiang Huang	47.92	–
Fu Shen	48.39	–
Lai Fu Zi	48.68	–
Nü Zhen Zi	48.80	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Mu Cao	49.23	–
Wu Mei	49.50	–
Huang Qi	49.69	–
Zhu Ru	50.03	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dan Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62573	62573	0.00	4.63
62574	62574	0.00	8.09

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dan Zhu Ye**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60147-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dan Zhu Ye; Lophatheri gracilis herba

Special notes

When selecting the *Dan Zhu Ye* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dan Zhu Ye	2	0	1

Second-stage model

For differentiation of the substance/substance group *Dan Zhu Ye* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dan Zhu Ye*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dan Zhu Ye	G046H1105821	62581	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105821	62582	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105921	62903	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105921	62904	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Dan Zhu Ye*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Dan Zhu Ye*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dan Zhu Ye	G046H1105821	62581†	20
PhytoComm	Dan Zhu Ye	G046H1105821	62582†	20
PhytoComm	Dan Zhu Ye	G046H1105921	62903†	20
PhytoComm	Dan Zhu Ye	G046H1105921	62904†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Dan Zhu Ye*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Dan Zhu Ye	G046H1105321	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dan Zhu Ye* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dan Zhu Ye* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	2	0	2	853

The substance/substance group *Dan Zhu Ye* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.8140 % (> 99.2290 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dan Zhu Ye* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yin Chen Hao	8.78	–
Sha Ren	12.81	–
Sang Bai Pi	14.04	–
Bu Gu Zhi	14.81	–
Jing Jie	15.35	–
Gu Sui Bu	18.42	–
He Shou Wu	18.63	–
Xuan Fu Hua	19.58	–
Sang Ji Shend	22.40	–
Wu Zhu Yu	22.51	–
Du Zhong	22.56	–
Mang Xiao	23.95	–
He Huan Pi	24.61	–
Xian Mao	26.08	–
Wu Jia Pi	26.17	–
Yu Xing Cao	26.76	–
Chen Pi	30.18	–
Nü Zhen Zi	30.82	–
Tu Si Zi	31.30	–
Jin Qian Cao	31.74	–
(Sheng) Di Huang	32.37	–
Ge Gen	35.11	–
Jiang Huang	37.98	–
Gou Teng	38.66	–
Yi Mu Cao	39.17	–
Sang Ye	40.33	–
Ding Xiang	42.06	–
(Bai) Dou Kou	42.52	–
Hong Jing Tian	43.94	–
(Shi) Chang Pu	44.42	–
Xuan Shen	45.87	–
Zhi Shi	46.98	–
Hua Shi	47.55	–
Ji Li	48.21	–
Wu Yao	48.38	–
Xin Yi	48.39	–
Xiao Hui Xiang	48.84	–
Han Lian Cao	49.13	–
Bai Hua She She Cao	49.65	–
Shan Yao	50.36	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dan Zhu Ye* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62581	62581	0.00	8.78
62582	62582	0.00	9.13
62903	62903	0.00	15.53
62904	62904	0.00	15.35

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dang Gui**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60003-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dang Gui; Angelicae sinensis radix

Special notes

When selecting the *Dang Gui* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dang Gui	4	0	4

Second-stage model

For differentiation of the substance/substance group *Dang Gui* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dang Gui*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dang Gui	G022H1306822	62413	40	from supplier
PhytoComm	Dang Gui	G022H1306822	62414	40	from supplier
PhytoComm	Dang Gui	G022HS290SK1	62649	40	from supplier
PhytoComm	Dang Gui	G022HS290SK1	62650	40	from supplier
PhytoComm	Dang Gui	G022H1306921	62963	40	from supplier
PhytoComm	Dang Gui	G022H1306921	62964	40	from supplier
PhytoComm	Dang Gui	G022HS290TH1	62965	40	from supplier
PhytoComm	Dang Gui	G022HS290TH1	62966	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 320 spectra of 8 reference samples from the substance/substance group *Dang Gui*. These samples are listed above in the *calibration samples* section. The reference samples come from 4 different batches.
- 24 280 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 8 reference samples from the substance/substance group *Dang Gui*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dang Gui	G022H1306822	62413†	20
PhytoComm	Dang Gui	G022H1306822	62414†	20
PhytoComm	Dang Gui	G022HS290SK1	62649†	20
PhytoComm	Dang Gui	G022HS290SK1	62650†	20

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dang Gui	G022H1306921	62963†	20
PhytoComm	Dang Gui	G022H1306921	62964†	20
PhytoComm	Dang Gui	G022HS290TH1	62965†	20
PhytoComm	Dang Gui	G022HS290TH1	62966†	20

- 12 317 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 12 spectra from 5 *Apo-Ident* customers from 5 batches from the substance/substance group *Dang Gui*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Dang Gui	g022h1306121	3
PhytoComm	Dang Gui	G022H1306121	1
Phytocomm	Dang Gui	G022H1306422	3
PhytoComm	Dang Gui	G022H1306621	4
PhytoComm	Dang Gui	G13-0784	1

- 845 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dang Gui* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dang Gui* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	320	0	24 280
Type B	0	152	8	12 317
Type C	2	5	7	843

†These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

The substance/substance group *Dang Gui* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 98.1250 %)
Type B	100.0000 % (> 99.9402 %)	95.0000 % (> 93.1250 %)
Type C	99.8450 % (> 99.2573 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dang Gui* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jie Geng	4.53	–
Cang Zhu	5.00	–
Dang Gui Wei	6.64	–
Gou Qi Zi	7.59	–
Yu Zhu	7.90	–
Chen Pi	10.17	–
Sang Zhi	11.03	–
Chuan Niu Xi	11.12	–
Gua Lou	11.20	–
San Qi	11.53	–
Du Huo	11.89	–
Long Yan Rou	11.90	–
Yuan Zhi	12.33	–
Zhi Mu	12.92	–
Ren Shen	13.10	–
Ban Lan Gen	14.03	–
Lian Qiao	14.59	–
Bai Shao Yao	14.72	–
Jin Yin Hua	14.76	–
Tian Hua Fen	14.85	–
Tai Zi Shen	14.89	–
Pi Pa Ye	15.34	–
Chuan Xiong	15.51	–
Mu Zei	16.02	–
Zhe Bei Mu	16.21	–
Ye Jiao Teng	16.22	–
Ze Xie	16.49	–
Ku Shen	16.58	–
Shan Yu Rou	16.59	–
Ren Dong Teng	17.13	–
Jiao Gu Lan	17.23	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ju Hua	17.37	–
Di Gu Pi	17.46	–
Hong Jing Tian	17.75	–
Bing Lang	18.00	–
Mu Gua	18.26	–
Fo Shou	18.94	–
Shan Yao	19.02	–
Dan Dou Chi	19.09	–
Chuan Lian Zi	19.15	–
He Huan Pi	19.26	–
Xiang Fu	19.53	–
Zhi Ke	19.64	–
Huang Qi	19.66	–
Ji Li	19.77	–
(Shi) Chang Pu	19.86	–
Ma Huang	20.01	–
Mu Dan Pi	20.34	–
Shen Qu	20.39	–
Chuang Mu Xiang	20.85	–
Ling Zhi	20.98	–
(Huai) Niu Xi	21.07	–
Lai Fu Zi	21.20	–
Ci Wu Jia	21.48	–
Suan Zao Ren	21.49	–
Tao Ren	21.56	–
Ba Ji Tian	21.92	–
Xie Bai	22.21	–
Fu Pen Zi	22.29	–
Qin Jiao	22.73	–
Bai Zhi	22.82	–
Bai Zi Ren	22.90	–
Lian Zi	23.13	–
Lu Gen	23.28	–
Hou Po	23.34	–
Gan Cao	23.39	–
Gou Teng	23.94	–
Fu Zi	23.98	–
Zhu Ru	23.99	–
E Zhu	24.33	–
Long Dan (Cao)	24.45	–
Sha Shen (Bei)	24.62	–
Fang Feng	24.77	–
Huang Qin	24.97	–
Rou Gui	25.14	–
Zhi Gan Cao	25.36	–
Dan Shen	25.55	–
Mi Huan Jun	25.76	–
Guang Huo Xiang	26.03	–
Gui Zhi	26.26	–
Wu Wei Zi	26.30	–
Huo Ma Ren	26.48	–
Gu Sui Bu	26.49	–
Ji Xue Teng	26.50	–
Wu Yao	26.61	–
Da Zao	26.75	–
Chuan Mu Tong	26.76	–
Mang Xiao	26.83	–
Ban Xia (Jiang)	26.85	–
Sheng Jiang	27.27	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	27.39	–
Yin Yang Huo	27.98	–
Fu Ling	28.60	–
Bai Zhu	28.81	–
Mai Ya	29.70	–
Bai He	29.74	–
Sha Ren	30.66	–
Yan Hu Suo	30.87	–
Jiang Huang	31.42	–
Bai Xian Pi	31.45	–
Zi Su Zi	31.48	–
Che Qian Zi	31.78	–
Zi Hua Di Ding	31.80	–
Mao Dong Qing	32.60	–
Chai Hu	32.65	–
Fu Xiao Mai	33.07	–
Mai Men Dong	34.07	–
Yi Yi Ren	34.93	–
(Fen) Bi Xie	35.14	–
Qiang Huo	35.41	–
She Gan	35.41	–
Ce Bai Ye	35.80	–
Chi Shao (Yao)	36.39	–
Sang Ji Shend	36.77	–
Hong Hua	36.83	–
Ma Huang Gen	37.21	–
Yu Jin	38.05	–
Bai Hua She She Cao	38.52	–
He Shou Wu	38.63	–
(Bai) Dou Kou	39.70	–
Ban Zhi Lian	39.94	–
Gan Jiang	40.16	–
Bo He	40.40	–
Qing Pi	40.81	–
Fu Shen	41.89	–
Zhu Ling	42.10	–
Jing Jie	43.15	–
Xu Duan	43.42	–
Huang Lian	43.50	–
Xiao Hui Xiang	44.31	–
Huang Bai	45.35	–
Niu Bang Zi	46.26	–
Cang Er Zi	46.89	–
Xin Yi	47.34	–
Du Zhong	47.93	–
Ge Gen	49.30	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dang Gui* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62413	62413	0.00	11.03
62414	62414	0.00	11.30
62649	62649	0.00	4.53
62650	62650	0.00	5.16
62963	62963	0.00	11.90
62964	62964	0.00	11.12
62965	62965	0.00	5.00
62966	62966	0.00	5.36

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dang Gui Wei**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60308-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dang Gui Wei; Angelicae sinensis radix extremitas

Special notes

When selecting the *Dang Gui Wei* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dang Gui Wei	1	0	0

Second-stage model

For differentiation of the substance/substance group *Dang Gui Wei* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dang Gui Wei*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dang Gui Wei	G312HS291SK1	62621	40	from supplier
PhytoComm	Dang Gui Wei	G312HS291SK1	62622	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Dang Gui Wei*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Dang Gui Wei*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dang Gui Wei	G312HS291SK1	62621†	20
PhytoComm	Dang Gui Wei	G312HS291SK1	62622†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Dang Gui Wei*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dang Gui Wei* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dang Gui Wei* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	3	39	1	12 434
Type C	0	0	0	857

The substance/substance group *Dang Gui Wei* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9821 % (> 99.9524 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dang Gui Wei* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Jie Geng	6.29	–
Yuan Zhi	7.56	–
Bai Shao Yao	7.85	–
Gua Lou	8.11	–
Cang Zhu	8.28	–
Yu Zhu	9.14	–
Dang Gui	9.14	–
Zhi Mu	9.61	–
Pi Pa Ye	10.44	–
Fo Shou	10.60	–
San Qi	11.04	–
Chen Pi	11.84	–
Chuan Xiong	12.39	–
Shen Qu	13.69	–
Tian Hua Fen	13.73	–
Tai Zi Shen	13.77	–
Ye Jiao Teng	13.95	–
Zhe Bei Mu	13.99	–
Mu Gua	14.09	–
Gou Qi Zi	14.10	–
Mu Zei	14.78	–
Di Gu Pi	14.91	–
Ban Lan Gen	15.21	–
Shan Yao	15.44	–
Ji Li	15.47	–
Ren Dong Teng	15.57	–
Ku Shen	16.13	–
Jin Yin Hua	16.43	–
Bai Zi Ren	16.58	–
Ren Shen	16.63	–
Lai Fu Zi	16.72	–
Lian Qiao	16.88	–
He Huan Pi	17.32	–
Long Yan Rou	17.58	–
Lian Zi	17.72	–
Ze Xie	18.33	–
Ling Zhi	18.73	–
Ci Wu Jia	18.81	–
Zhu Ru	19.20	–
Xie Bai	19.92	–
Rou Gui	20.15	–
Sheng Jiang	20.55	–
Tu Fu Ling	20.76	–
Tao Ren	20.77	–
Jiao Gu Lan	20.79	–
Suan Zao Ren	20.98	–
Dan Dou Chi	21.38	–
Huo Ma Ren	21.47	–
Shan Yu Rou	21.63	–
Chuang Mu Xiang	22.95	–
Fu Zi	23.13	–
Lu Gen	23.24	–
Gou Teng	23.31	–
Fu Ling	23.48	–
Fu Pen Zi	23.57	–
Gui Zhi	23.61	–
Hong Jing Tian	24.07	–
Zhi Ke	24.23	–
Ban Xia (Jiang)	24.53	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gu Sui Bu	25.38	–
Hou Po	25.54	–
Fu Xiao Mai	25.71	–
Dan Shen	26.03	–
Ji Xue Teng	26.87	–
Ma Huang	26.90	–
Bai Xian Pi	26.95	–
Wu Wei Zi	27.88	–
Mao Dong Qing	27.94	–
Gan Cao	28.64	–
Huang Qin	28.69	–
Yi Yi Ren	30.28	–
Che Qian Zi	30.63	–
Chai Hu	30.96	–
Yin Yang Huo	31.51	–
Mang Xiao	31.94	–
Guang Huo Xiang	32.14	–
(Fen) Bi Xie	32.93	–
Ma Huang Gen	33.00	–
Ce Bai Ye	33.40	–
Yan Hu Suo	34.10	–
She Gan	35.18	–
Yu Jin	35.40	–
Sang Zhi	36.42	–
Mai Men Dong	36.76	–
Qiang Huo	36.92	–
Zhu Ling	38.60	–
Chi Shao (Yao)	38.61	–
Zhi Gan Cao	38.92	–
Fu Shen	39.89	–
Zi Hua Di Ding	41.50	–
Qing Pi	42.82	–
Bo He	43.52	–
Jing Jie	45.17	–
Hong Hua	45.38	–
Huang Lian	45.71	–
Chuan Lian Zi	47.43	–
Cang Er Zi	48.17	–
Sha Ren	48.33	–
Huang Qi	49.07	–
Huang Bai	49.22	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dang Gui Wei* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62621	62621	0.00	6.36
62622	62622	0.00	6.29

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Di Fu Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002264-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Di Fu Zi; Kochiae scopariae fructus

Special notes

When selecting the *Di Fu Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Di Fu Zi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Di Fu Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Di Fu Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Di Fu Zi	G131H0605921	62737	40	from supplier
PhytoComm	Di Fu Zi	G131H0605921	62738	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Di Fu Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Di Fu Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Di Fu Zi	G131H0605921	62737†	20
PhytoComm	Di Fu Zi	G131H0605921	62738†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Di Fu Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Di Fu Zi	G131H0605122	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Di Fu Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Di Fu Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Di Fu Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Di Fu Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	23.68	–
Guang Huo Xiang	32.70	–
Hua Shi	46.19	–
Wu Mei	54.59	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Di Fu Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62737	62737	0.00	23.70
62738	62738	0.00	23.68

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Di Gu Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60036-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Di Gu Pi; Lycii chinensis radicis cortex

Special notes

When selecting the *Di Gu Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Di Gu Pi	3	0	2

Second-stage model

For differentiation of the substance/substance group *Di Gu Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Di Gu Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Di Gu Pi	G149H0601821	62347	40	from supplier
PhytoComm	Di Gu Pi	G149H0601821	62348	40	from supplier
PhytoComm	Di Gu Pi	G149HS101SK1	62691	40	from supplier
PhytoComm	Di Gu Pi	G149HS101SK1	62692	40	from supplier
PhytoComm	Di Gu Pi	G149H0601922	62997	40	from supplier
PhytoComm	Di Gu Pi	G149H0601922	62998	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Di Gu Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Di Gu Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Di Gu Pi	G149H0601821	62347†	20
PhytoComm	Di Gu Pi	G149H0601821	62348†	20
PhytoComm	Di Gu Pi	G149HS101SK1	62691†	20
PhytoComm	Di Gu Pi	G149HS101SK1	62692†	20
PhytoComm	Di Gu Pi	G149H0601922	62997†	20
PhytoComm	Di Gu Pi	G149H0601922	62998†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Di Gu Pi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Di Gu Pi	g149h060112	1
Phytocomm	Di Gu Pi	G149H0601421	1
PhytoComm	Di Gu Pi	G149H0601421	2

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Di Gu Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Di Gu Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	238	2	24 360
Type B	7	118	2	12 350
Type C	3	1	3	850

The substance/substance group *Di Gu Pi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	99.1667 % (> 97.9167 %)
Type B	99.9429 % (> 99.9130 %)	98.3333 % (> 95.8333 %)
Type C	99.6589 % (> 99.0724 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Di Gu Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sheng Jiang	4.73	–
He Huan Pi	4.80	–
Gui Zhi	5.65	–
Sha Shen (Bei)	5.80	–
Ci Wu Jia	5.91	–
Zhu Ru	6.95	–
Bai Shao Yao	7.59	–
Ban Xia (Jiang)	8.02	–
Fu Xiao Mai	9.04	–
Sang Zhi	9.34	–
Lai Fu Zi	9.59	–
Tao Ren	9.70	–
Bai Zi Ren	10.42	–
Yi Yi Ren	10.60	–
Ji Li	10.80	–
Rou Gui	10.83	–
Zhe Bei Mu	10.91	–
Shen Qu	11.07	–
Lian Zi	11.77	–
Gan Jiang	11.96	–
Fu Ling	12.98	–
Tai Zi Shen	13.56	–
Chuan Lian Zi	13.82	–
Ye Jiao Teng	14.13	–
Gou Teng	14.27	–
Mi Huan Jun	14.43	–
E Zhu	14.45	–
Huo Ma Ren	14.81	–
Fo Shou	14.82	–
Tian Hua Fen	15.02	–
Ling Zhi	15.11	–
Mu Zei	15.19	–
Yan Hu Suo	15.27	–
Gua Lou	15.75	–
Bai Xian Pi	15.88	–
Ji Xue Teng	16.85	–
Jie Geng	17.53	–
Dang Gui	17.64	–
Chuan Mu Tong	17.66	–
Zi Su Zi	17.90	–
Yuan Zhi	19.07	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pi Pa Ye	19.41	–
Shan Yao	19.42	–
Gu Sui Bu	19.70	–
Jiang Huang	20.32	–
(Huai) Niu Xi	20.99	–
Ze Xie	21.02	–
Ma Huang Gen	21.25	–
Lu Gen	21.51	–
Mai Ya	21.52	–
Cang Er Zi	22.15	–
Long Dan (Cao)	22.72	–
Lian Qiao	22.75	–
Ren Dong Teng	23.10	–
Tu Fu Ling	23.28	–
Yu Jin	23.35	–
(Bai) Dou Kou	24.05	–
Chi Shao (Yao)	24.07	–
Chuan Xiong	24.11	–
Suan Zao Ren	24.13	–
Mang Xiao	24.29	–
Fu Shen	24.56	–
Bai Zhu	24.83	–
(Shi) Chang Pu	25.57	–
Jin Yin Hua	25.81	–
Sha Ren	25.92	–
Qiang Huo	25.96	–
Niu Bang Zi	26.06	–
Huang Qi	26.48	–
Zhi Gan Cao	26.69	–
Zhu Ling	26.75	–
Chen Pi	27.17	–
Bai He	27.23	–
Zhi Mu	27.34	–
Mu Gua	27.44	–
Wu Wei Zi	28.89	–
Cang Zhu	29.33	–
Fu Zi	29.45	–
Chuan Niu Xi	30.11	–
(Fen) Bi Xie	30.37	–
Ce Bai Ye	30.46	–
Fu Pen Zi	31.14	–
Shan Yu Rou	31.17	–
Chai Hu	31.24	–
Jiao Gu Lan	31.43	–
Qin Jiao	31.72	–
Ren Shen	31.82	–
Ban Zhi Lian	32.27	–
Dan Dou Chi	32.35	–
Ku Shen	32.66	–
San Qi	32.74	–
Dang Gui Wei	32.75	–
Bo He	32.87	–
Dan Shen	32.88	–
Zhi Ke	32.92	–
She Gan	33.82	–
Bai Zhi	34.06	–
Yu Zhu	34.19	–
Gan Cao	34.69	–
Da Zao	34.72	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	35.26	–
Xiang Fu	35.66	–
Hou Po	35.82	–
Huang Qin	36.01	–
Ban Lan Gen	36.08	–
Mao Dong Qing	36.18	–
Yin Yang Huo	36.60	–
Pu Gong Ying	37.20	–
Gou Qi Zi	37.46	–
Xiao Hui Xiang	38.22	–
Chuang Mu Xiang	38.54	–
Che Qian Zi	39.34	–
Xie Bai	39.78	–
Guang Huo Xiang	40.31	–
Bai Hua She She Cao	40.79	–
Huang Lian	42.08	–
Ma Huang	42.31	–
Jing Jie	42.55	–
Du Huo	42.91	–
Huang Bai	42.93	–
Du Zhong	43.43	–
Qing Pi	44.84	–
Hong Jing Tian	45.45	–
Yi Mu Cao	45.76	–
Dong Gua Zi	45.90	–
Mu Dan Pi	46.69	–
Ba Ji Tian	47.53	–
Sang Ji Shend	47.62	–
Wu Yao	49.77	–
Long Yan Rou	49.79	–
Sang Ye	49.90	–
Ju Hua	50.04	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Di Gu Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62347	62347	0.00	9.34
62348	62348	0.00	9.72
62691	62691	0.00	5.09
62692	62692	0.00	4.73
62997	62997	0.00	5.80
62998	62998	0.00	6.45

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ding Xiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50309-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ding Xiang; Caryophylli flos

Special notes

When selecting the *Ding Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ding Xiang	1	0	0

Second-stage model

For differentiation of the substance/substance group *Ding Xiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ding Xiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ding Xiang	G058H0210021	63013	40	from supplier
PhytoComm	Ding Xiang	G058H0210021	63014	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ding Xiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ding Xiang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ding Xiang	G058H0210021	63013†	20
PhytoComm	Ding Xiang	G058H0210021	63014†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Ding Xiang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ding Xiang	G058H0210021	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ding Xiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ding Xiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Ding Xiang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ding Xiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	18.73	–
(Sheng) Di Huang	24.06	–
Yu Xing Cao	24.72	–
Xian Mao	24.87	–
Jin Qian Cao	27.56	–
Jing Jie	30.69	–
Du Zhong	34.30	–
Sang Ye	34.40	–
Ge Gen	36.35	–
Shu Di (Huang)	38.35	–
Wu Jia Pi	39.36	–
Dan Zhu Ye	40.04	–
Sang Ji Shend	41.87	–
Xuan Shen	44.52	–
Hua Shi	47.28	–
Guang Huo Xiang	47.78	–
Gou Teng	50.04	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ding Xiang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
63013	63013	0.00	18.84
63014	63014	0.00	18.73

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Dong Gua Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60202-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Dong Gua Zi; Benincasae hispidae semen

Special notes

When selecting the *Dong Gua Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Dong Gua Zi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Dong Gua Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Dong Gua Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dong Gua Zi	G049H0551822	62513	40	from supplier
PhytoComm	Dong Gua Zi	G049H0551822	62514	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Dong Gua Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Dong Gua Zi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Dong Gua Zi	G049H0551822	62513†	20
PhytoComm	Dong Gua Zi	G049H0551822	62514†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Dong Gua Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Dong Gua Zi	G049H0551221	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Dong Gua Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Dong Gua Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Dong Gua Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Dong Gua Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ku Shen	29.73	–
Shan Yao	30.81	–
Shen Qu	34.98	–
Mang Xiao	37.56	–
Di Gu Pi	47.16	–
Cang Er Zi	48.77	–
Bai Shao Yao	49.23	–
Chi Shao (Yao)	49.50	–
Sang Zhi	49.85	–
Wu Wei Zi	50.71	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dong Gua Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62513	62513	0.00	29.73
62514	62514	0.00	29.90

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Du Huo**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10005053-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Du Huo; Angelicae pubescentis radix

Special notes

When selecting the *Du Huo* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Du Huo	1	0	2

Second-stage model

For differentiation of the substance/substance group *Du Huo* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Du Huo*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Du Huo	G021H1610921	63021	40	from supplier
PhytoComm	Du Huo	G021H1610921	63022	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Du Huo*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Du Huo*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Du Huo	G021H1610921	63021†	20
PhytoComm	Du Huo	G021H1610921	63022†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Du Huo*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Du Huo	G021H1610421	2
Phytocomm	Du Huo	g021h610221	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Du Huo* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Du Huo* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	3	854

The substance/substance group *Du Huo* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Du Huo* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui	9.05	–
Chuan Niu Xi	10.50	–
Bing Lang	18.27	–
(Shi) Chang Pu	18.36	–
Chuan Lian Zi	18.81	–
Sang Zhi	19.80	–
Huang Qi	20.04	–
Shan Yu Rou	20.63	–
Bai Zhi	21.32	–
Ju Hua	21.43	–
Chen Pi	22.04	–
Xiang Fu	22.72	–
Ji Li	23.58	–
Zhi Gan Cao	25.12	–
Tian Hua Fen	25.49	–
(Huai) Niu Xi	26.11	–
Mu Dan Pi	26.65	–
Mang Xiao	27.14	–
Fang Feng	27.68	–
Ba Ji Tian	27.83	–
E Zhu	28.07	–
Jie Geng	28.67	–
Zi Su Zi	30.81	–
Bai Zhu	31.34	–
Di Gu Pi	31.42	–
Mi Huan Jun	31.72	–
Wu Yao	32.74	–
Sha Ren	32.93	–
He Huan Pi	33.01	–
Shan Yao	33.04	–
Jiang Huang	33.42	–
Lian Qiao	33.71	–
Qin Jiao	33.73	–
Gua Lou	33.76	–
Da Zao	34.73	–
Yan Hu Suo	35.50	–
Zhi Ke	36.24	–
Sha Shen (Bei)	36.62	–
Lai Fu Zi	36.63	–
Chi Shao (Yao)	37.62	–
Bai Hua She She Cao	38.40	–
Gu Sui Bu	38.56	–
(Bai) Dou Kou	38.88	–
Ku Shen	40.02	–
Long Dan (Cao)	40.03	–
Yuan Zhi	42.10	–
Gan Jiang	42.20	–
Sang Ji Shend	42.36	–
Xu Duan	43.08	–
Chuan Mu Tong	44.30	–
Xiao Hui Xiang	44.62	–
Niu Bang Zi	44.94	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Wei Zi	45.33	–
He Shou Wu	45.93	–
Mai Ya	46.02	–
Bai He	47.56	–
Hong Jing Tian	47.66	–
Cang Er Zi	47.67	–
Mao Dong Qing	48.54	–
Pu Gong Ying	48.59	–
Du Zhong	48.68	–
Tu Fu Ling	49.29	–
Jin Yin Hua	49.47	–
Xin Yi	49.67	–
Huang Lian	49.97	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Du Huo* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
63021	63021	0.00	9.05
63022	63022	0.00	9.08

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Du Zhong**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60214-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Du Zhong; Eucommiae ulmoidis cortex

Special notes

When selecting the *Du Zhong* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Du Zhong	3	0	4

Second-stage model

For differentiation of the substance/substance group *Du Zhong* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Du Zhong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Du Zhong	G101H0736821	62335	40	from supplier
PhytoComm	Du Zhong	G101H0736821	62336	40	from supplier
PhytoComm	Du Zhong	G101HS133SH1	62579	40	from supplier
PhytoComm	Du Zhong	G101HS133SH1	62580	40	from supplier
PhytoComm	Du Zhong	G101HS133SP1	62729	40	from supplier
PhytoComm	Du Zhong	G101HS133SP1	62730	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Du Zhong*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Du Zhong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Du Zhong	G101H0736821	62335†	20
PhytoComm	Du Zhong	G101H0736821	62336†	20
PhytoComm	Du Zhong	G101HS133SH1	62579†	20
PhytoComm	Du Zhong	G101HS133SH1	62580†	20
PhytoComm	Du Zhong	G101HS133SP1	62729†	20
PhytoComm	Du Zhong	G101HS133SP1	62730†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Du Zhong*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Du Zhong	g101h0736021	1
Phytocomm	Du Zhong	G101H0736321	1
Phytocomm	Du Zhong	G101H0736522	5
Phytocomm	Du Zhong	g101h076521	1

- 849 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Du Zhong* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Du Zhong* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	1	118	2	12 356
Type C	0	1	7	849

The substance/substance group *Du Zhong* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	99.9929 % (> 99.9630 %)	98.3333 % (> 95.8333 %)
Type C	100.0000 % (> 98.8253 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Du Zhong* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Qing Hao	4.20	–
Mao Dong Qing	4.79	–
Yu Jin	5.09	–
Qiang Huo	7.49	–
Chai Hu	7.78	–
Sha Ren	7.89	–
Ling Zhi	8.14	–
Bo He	8.24	–
Hou Po	9.11	–
Jing Jie	9.17	–
Bai Jiang Cao	9.91	–
Zi Hua Di Ding	10.10	–
Bai Xian Pi	11.24	–
Yin Yang Huo	11.29	–
Fu Zi	11.68	–
Dan Dou Chi	12.67	–
Qing Pi	13.38	–
Yan Hu Suo	13.80	–
Jiao Gu Lan	14.26	–
Ye Jiao Teng	14.76	–
Ce Bai Ye	15.09	–
Shen Qu	15.10	–
Shan Yao	15.56	–
Dan Shen	16.57	–
Che Qian Zi	16.68	–
Fu Pen Zi	16.97	–
Pi Pa Ye	17.10	–
Ji Li	17.67	–
Hong Jing Tian	18.45	–
Ren Dong Teng	18.49	–
Tu Fu Ling	18.77	–
Chuan Xiong	18.78	–
Yin Chen Hao	18.80	–
Tian Hua Fen	19.13	–
Huang Bai	19.15	–
Zhu Ling	19.22	–
Dan Zhu Ye	19.33	–
Lian Zi	19.65	–
Wu Jia Pi	19.66	–
Guang Huo Xiang	19.91	–
Huang Lian	21.17	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang	21.44	–
Zhi Ke	21.91	–
Ji Xue Teng	22.02	–
Pu Gong Ying	22.03	–
Gu Sui Bu	22.24	–
Jin Yin Hua	23.29	–
Wu Wei Zi	23.36	–
Ze Lan	24.24	–
Lian Qiao	24.27	–
Mang Xiao	24.48	–
Xi Xian Cao	24.53	–
Di Gu Pi	25.25	–
He Huan Pi	26.02	–
Ban Lan Gen	26.07	–
Bai Shao Yao	26.23	–
Sang Ji Shend	26.29	–
Bu Gu Zhi	27.05	–
Gua Lou	27.28	–
Ren Shen	27.60	–
Nü Zhen Zi	27.82	–
Gou Teng	27.95	–
Suan Zao Ren	28.31	–
Yuan Zhi	29.36	–
Ma Huang Gen	29.59	–
(Fen) Bi Xie	29.92	–
He Shou Wu	29.93	–
Xuan Fu Hua	30.23	–
Zhe Bei Mu	30.39	–
Sheng Jiang	30.47	–
Gan Cao	30.99	–
Jin Qian Cao	31.04	–
Rou Gui	31.35	–
Cang Zhu	31.52	–
Sang Ye	31.73	–
Xuan Shen	31.89	–
Gou Qi Zi	32.14	–
She Gan	32.51	–
Chuang Mu Xiang	32.53	–
Ding Xiang	32.80	–
Mu Zei	32.94	–
Yi Mu Cao	33.07	–
Fu Ling	33.35	–
Lu Gen	33.42	–
Sang Bai Pi	33.88	–
Yi Yi Ren	35.58	–
Hong Hua	35.62	–
Zhi Gan Cao	35.66	–
E Zhu	35.68	–
Wu Zhu Yu	35.72	–
Cang Er Zi	35.84	–
Huo Ma Ren	35.85	–
Gui Zhi	36.84	–
Jiang Huang	37.48	–
Zhi Shi	38.25	–
Ban Zhi Lian	38.92	–
Chi Shao (Yao)	39.48	–
Yu Xing Cao	39.52	–
Tao Ren	39.87	–
Wu Mei	39.88	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Zhi	40.28	–
Ge Gen	41.17	–
Xin Yi	41.26	–
Chen Pi	41.69	–
Jie Geng	42.61	–
Tai Zi Shen	42.93	–
Xie Bai	43.15	–
Shan Yu Rou	43.39	–
Bai Zi Ren	43.40	–
Huang Qin	44.13	–
Ze Xie	44.86	–
Zhu Ru	45.06	–
Niu Bang Zi	45.35	–
Ban Xia (Jiang)	46.21	–
Ci Wu Jia	47.02	–
Yu Zhu	47.99	–
Xia Ku Cao	48.47	–
Dang Gui	48.70	–
Fu Xiao Mai	49.45	–
San Qi	49.58	–
Hua Shi	49.58	–
Fu Shen	49.80	–
Sha Shen (Bei)	49.90	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Du Zhong* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62335	62335	0.00	12.52
62336	62336	0.00	13.01
62579	62579	0.00	4.79
62580	62580	0.00	4.98
62729	62729	0.00	4.20
62730	62730	0.00	5.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**E Zhu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50344-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

E Zhu; Curcumae zedoariae rhizoma

Special notes

When selecting the *E Zhu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
E Zhu	1	0	2

Second-stage model

For differentiation of the substance/substance group *E Zhu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *E Zhu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	E Zhu	G083H1110821	62515	40	from supplier
PhytoComm	E Zhu	G083H1110821	62516	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *E Zhu*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *E Zhu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	E Zhu	G083H1110821	62515†	20
PhytoComm	E Zhu	G083H1110821	62516†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *E Zhu*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	E Zhu	G083H1110222	2
Phytocomm	E Zhu	h1110021	1
Phytocomm	E Zhu	H1110021	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *E Zhu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *E Zhu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	2	0	4	851

The substance/substance group *E Zhu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8062 % (> 99.2197 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *E Zhu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Li	15.45	–
Jiang Huang	17.25	–
Cang Er Zi	17.37	–
Sha Shen (Bei)	17.91	–
Gan Jiang	19.17	–
(Bai) Dou Kou	22.11	–
Di Gu Pi	22.13	–
Mang Xiao	23.29	–
Yan Hu Suo	23.44	–
Sha Ren	23.97	–
Zi Su Zi	24.69	–
He Huan Pi	24.83	–
Niu Bang Zi	26.51	–
Lai Fu Zi	27.84	–
Chuan Lian Zi	28.13	–
Sang Zhi	29.14	–
Shan Yao	32.50	–
Chen Pi	33.25	–
Mi Huan Jun	35.34	–
Suan Zao Ren	38.75	–
Mai Ya	39.53	–
Chuan Mu Tong	39.65	–
Dang Gui	39.66	–
Bai Hua She She Cao	39.93	–
(Huai) Niu Xi	40.42	–
Fu Ling	41.35	–
Long Dan (Cao)	42.17	–
Rou Gui	42.25	–
Huang Lian	42.32	–
Du Huo	42.78	–
Ban Zhi Lian	43.37	–
Wang Bu Liu Xing	43.84	–
Tian Hua Fen	44.15	–
Lian Qiao	44.19	–
(Shi) Chang Pu	44.34	–
Du Zhong	44.58	–
Da Zao	44.78	–
Chuan Niu Xi	46.80	–
Xi Xian Cao	46.92	–
Zhi Ke	47.66	–
Huang Qi	47.84	–
Qiang Huo	48.08	–
Xiang Fu	48.64	–
Lian Zi	49.29	–
Xin Yi	49.61	–
Dan Dou Chi	50.44	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *E Zhu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62515	62515	0.00	15.45
62516	62516	0.00	15.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fang Feng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	`50261-2022-02-15`
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fang Feng; Saposhnikoviae radix

Special notes

When selecting the *Fang Feng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fang Feng	1	0	3

Second-stage model

For differentiation of the substance/substance group *Fang Feng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fang Feng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fang Feng	G217H0728922	62859	40	from supplier
PhytoComm	Fang Feng	G217H0728922	62860	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Fang Feng*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Fang Feng*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fang Feng	G217H0728922	62859†	20
PhytoComm	Fang Feng	G217H0728922	62860†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 7 spectra from 5 *Apo-Ident* customers from 3 batches from the substance/substance group *Fang Feng*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Fang Feng	g217h0728221	1
Phytocomm	Fang Feng	G217H0728521	3
PhytoComm	Fang Feng	G217H0728521	2
Phytocomm	Fang Feng	g217h0728923	1

- 850 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fang Feng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fang Feng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	1	0	7	849

The substance/substance group *Fang Feng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.7674 % (> 99.1802 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fang Feng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chen Pi	17.02	–
Ju Hua	19.77	–
Hong Jing Tian	21.56	–
Mang Xiao	25.68	–
Chuan Lian Zi	28.50	–
Shan Yu Rou	28.69	–
Chuan Niu Xi	28.75	–
Gu Sui Bu	29.07	–
(Shi) Chang Pu	30.14	–
Sang Bai Pi	30.22	–
Sha Ren	30.88	–
Gua Lou	30.99	–
Sang Ji Shend	31.31	–
Dang Gui	31.62	–
Xiang Fu	32.12	–
Yi Mu Cao	33.18	–
Wu Zhu Yu	33.25	–
He Huan Pi	34.14	–
Huang Qi	35.63	–
Ku Shen	36.19	–
Jiang Huang	36.71	–
Ba Ji Tian	37.08	–
Wu Yao	37.62	–
Mu Dan Pi	38.03	–
Bing Lang	38.79	–
Zhi Gan Cao	41.38	–
Ge Gen	42.06	–
Yin Chen Hao	42.33	–
Ji Li	43.35	–
He Shou Wu	44.23	–
Sang Zhi	44.67	–
Xu Duan	44.68	–
Dan Zhu Ye	45.01	–
Shan Yao	45.13	–
Chi Shao (Yao)	45.72	–
Xiao Hui Xiang	45.75	–
Cang Er Zi	45.86	–
Nü Zhen Zi	46.41	–
Xin Yi	48.25	–
E Zhu	49.14	–
Tian Hua Fen	49.57	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fang Feng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62859	62859	0.00	17.02
62860	62860	0.00	17.26

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fo Shou**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60119-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fo Shou; Citri sarcodactylis fructus

Special notes

When selecting the *Fo Shou* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fo Shou	1	0	0

Second-stage model

For differentiation of the substance/substance group *Fo Shou* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fo Shou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fo Shou	G307HS159RQ1	62419	40	from supplier
PhytoComm	Fo Shou	G307HS159RQ1	62420	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Fo Shou*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Fo Shou*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fo Shou	G307HS159RQ1	62419†	20
PhytoComm	Fo Shou	G307HS159RQ1	62420†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Fo Shou*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fo Shou* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fo Shou* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	79	1	24 519
Type B	0	39	1	12 437
Type C	0	0	0	857

The substance/substance group *Fo Shou* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9976 % (> 99.9827 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9406 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fo Shou* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Gua Lou	4.62	–
Ye Jiao Teng	9.33	–
Mu Zei	10.81	–
Bai Shao Yao	11.69	–
Tao Ren	11.97	–
Tai Zi Shen	12.17	–
Rou Gui	13.02	–
He Huan Pi	13.04	–
Ze Xie	13.18	–
Di Gu Pi	14.06	–
Shen Qu	14.20	–
Chen Pi	14.27	–
Yu Zhu	14.61	–
Lai Fu Zi	14.84	–
Jin Yin Hua	15.00	–
Lian Zi	15.00	–
Ban Xia (Jiang)	15.37	–
Ci Wu Jia	15.40	–
Zhe Bei Mu	15.51	–
Gui Zhi	15.86	–
Yuan Zhi	16.55	–
Ling Zhi	16.60	–
Zhi Mu	16.65	–
Bai Zi Ren	16.69	–
Cang Zhu	16.70	–
Jie Geng	16.70	–
Pi Pa Ye	16.76	–
Tian Hua Fen	16.83	–
Chuan Xiong	16.92	–
Ji Li	16.93	–
Sheng Jiang	17.11	–
Bai Xian Pi	18.00	–
Dang Gui Wei	18.10	–
Fu Xiao Mai	18.36	–
Lu Gen	18.69	–
Gu Sui Bu	18.97	–
Fu Ling	18.98	–
Gou Teng	19.07	–
Tu Fu Ling	19.51	–
Dang Gui	19.67	–
Ren Dong Teng	20.24	–
Zhu Ru	20.43	–
Ban Lan Gen	20.59	–
San Qi	20.70	–
Suan Zao Ren	21.07	–
Huo Ma Ren	21.41	–
Ji Xue Teng	22.67	–
Shan Yao	22.69	–
Ma Huang Gen	23.54	–
Mu Gua	24.07	–
Ren Shen	24.40	–
Lian Qiao	24.73	–
Yi Yi Ren	25.89	–
Long Yan Rou	26.11	–
Ma Huang	26.59	–
Fu Pen Zi	26.75	–
(Fen) Bi Xie	27.77	–
Fu Shen	27.82	–
Jiao Gu Lan	27.85	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yu Rou	28.19	–
Dan Shen	28.34	–
Fu Zi	28.52	–
Zhi Ke	28.62	–
Dan Dou Chi	29.04	–
She Gan	29.75	–
Hong Jing Tian	29.89	–
Gou Qi Zi	30.17	–
Huang Qin	31.13	–
Yu Jin	31.60	–
Ku Shen	31.64	–
Mang Xiao	31.89	–
Ce Bai Ye	32.41	–
Yin Yang Huo	32.45	–
Xie Bai	32.77	–
Chai Hu	33.11	–
Mao Dong Qing	33.29	–
Chuang Mu Xiang	33.59	–
Guang Huo Xiang	35.16	–
Hou Po	35.44	–
Yan Hu Suo	35.97	–
Zhu Ling	36.62	–
Wu Wei Zi	36.81	–
Sang Zhi	39.13	–
Gan Cao	39.77	–
Che Qian Zi	40.12	–
Qiang Huo	40.72	–
Jing Jie	43.25	–
Mai Men Dong	44.53	–
Chi Shao (Yao)	45.64	–
Chuan Lian Zi	45.82	–
Zi Hua Di Ding	46.26	–
Huang Lian	48.15	–
Bai Zhu	48.79	–
Bo He	49.16	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fo Shou* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62419	62419	0.00	6.62
62420	62420	0.00	4.62

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by

laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fu Ling**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50260-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fu Ling; Poriae cocos sclerotium

Special notes

When selecting the *Fu Ling* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fu Ling	3	0	4

Second-stage model

For differentiation of the substance/substance group *Fu Ling* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fu Ling*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fu Ling	G200HS211RT1	62429	40	from supplier
PhytoComm	Fu Ling	G200HS211RT1	62430	40	from supplier
PhytoComm	Fu Ling	G200HS211SH1	62597	40	from supplier
PhytoComm	Fu Ling	G200HS211SH1	62598	40	from supplier
PhytoComm	Fu Ling	G200HS211TK1	62849	40	from supplier
PhytoComm	Fu Ling	G200HS211TK1	62850	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Fu Ling*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Fu Ling*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fu Ling	G200HS211RT1	62429†	20
PhytoComm	Fu Ling	G200HS211RT1	62430†	20
PhytoComm	Fu Ling	G200HS211SH1	62597†	20
PhytoComm	Fu Ling	G200HS211SH1	62598†	20
PhytoComm	Fu Ling	G200HS211TK1	62849†	20
PhytoComm	Fu Ling	G200HS211TK1	62850†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 12 spectra from 7 *Apo-Ident* customers from 5 batches from the substance/substance group *Fu Ling*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phyto Comm	Fu Ling	G200H1019522	1
Phytocomm	Fu Ling	G200H1019221	1
PhytoComm	Fu Ling	G200H1019221	1
Phytocomm	Fu Ling	g200h1019321	3
Phytocomm	Fu Ling	G200H1019321	1
PhytoComm	Fu Ling	G200H1019321	1
Phytocomm	Fu Ling	G200H1019522	2
PhytoComm	Fu Ling	G200H1019522	1
PhytoComm	Fu Ling	G200H1019622	1

- 845 spectra from 12 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fu Ling* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fu Ling* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	119	1	12 357
Type C	1	0	12	844

The substance/substance group *Fu Ling* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	99.1667 % (> 96.6667 %)
Type C	99.7674 % (> 99.1798 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fu Ling* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yi Yi Ren	4.55	–
Gui Zhi	6.82	–
Fu Shen	7.02	–
Rou Gui	8.29	–
Ma Huang Gen	8.76	–
Sheng Jiang	8.85	–
He Huan Pi	9.15	–
Fu Xiao Mai	10.01	–
Gou Teng	10.53	–
Ji Xue Teng	12.57	–
Mu Zei	12.83	–
Lai Fu Zi	13.34	–
Zhu Ru	14.13	–
Ling Zhi	14.28	–
Ban Xia (Jiang)	14.97	–
Di Gu Pi	15.15	–
Bai Xian Pi	15.70	–
Lian Zi	15.80	–
Ci Wu Jia	16.46	–
Ji Li	16.78	–
Tao Ren	16.83	–
Ren Dong Teng	18.56	–
Huo Ma Ren	19.42	–
Zhu Ling	19.62	–
Gu Sui Bu	19.70	–
Lu Gen	20.32	–
Bai Zi Ren	20.73	–
She Gan	20.96	–
(Fen) Bi Xie	22.00	–
Fo Shou	22.25	–
Bai Shao Yao	22.32	–
Ye Jiao Teng	22.59	–
Ce Bai Ye	23.08	–
Tai Zi Shen	23.14	–
Shen Qu	23.95	–
Zhe Bei Mu	26.39	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Jin	26.46	–
Suan Zao Ren	26.56	–
Dan Dou Chi	26.76	–
Tu Fu Ling	28.45	–
Gua Lou	29.73	–
Chen Pi	30.10	–
Shan Yao	31.12	–
Mao Dong Qing	31.36	–
Yuan Zhi	31.48	–
Chai Hu	31.93	–
Tian Hua Fen	32.05	–
Ze Xie	32.18	–
Mang Xiao	32.77	–
Chuan Xiong	33.51	–
Dan Shen	35.50	–
Lian Qiao	35.94	–
Pi Pa Ye	36.19	–
Yan Hu Suo	36.43	–
Jiao Gu Lan	37.40	–
Yin Yang Huo	37.44	–
Dang Gui Wei	37.66	–
Ren Shen	39.90	–
Zhi Mu	40.41	–
Hong Jing Tian	40.46	–
Zhi Ke	41.18	–
Guang Huo Xiang	41.55	–
Yu Zhu	41.63	–
San Qi	42.06	–
Che Qian Zi	43.03	–
Ban Lan Gen	43.32	–
Jie Geng	43.49	–
Dang Gui	43.59	–
Shan Yu Rou	43.71	–
Cang Zhu	43.72	–
Qiang Huo	43.99	–
Jin Yin Hua	44.55	–
Fu Zi	45.02	–
Gan Cao	45.13	–
Wu Wei Zi	45.29	–
Huang Qin	45.40	–
Zi Hua Di Ding	46.16	–
Sang Zhi	46.17	–
Jing Jie	47.83	–
Sha Ren	48.77	–
Hou Po	48.91	–
Huang Lian	49.16	–
Ma Huang	49.38	–
Fu Pen Zi	49.57	–
Bai Zhu	50.15	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fu Ling* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62429	62429	0.00	9.51
62430	62430	0.00	8.75
62597	62597	0.00	9.39
62598	62598	0.00	9.04
62849	62849	0.00	4.55
62850	62850	0.00	5.32

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fu Pen Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60044-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fu Pen Zi; Rubi chingii fructus

Special notes

When selecting the *Fu Pen Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fu Pen Zi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Fu Pen Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fu Pen Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fu Pen Zi	G213HS343TH1	62855	40	from supplier
PhytoComm	Fu Pen Zi	G213HS343TH1	62856	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Fu Pen Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Fu Pen Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fu Pen Zi	G213HS343TH1	62855†	20
PhytoComm	Fu Pen Zi	G213HS343TH1	62856†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Fu Pen Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Fu Pen Zi	g213h1813122	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fu Pen Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fu Pen Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	39	1	12 437
Type C	0	0	1	856

The substance/substance group *Fu Pen Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fu Pen Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Fu Zi	5.93	–
Hou Po	7.64	–
Yan Hu Suo	7.91	–
Guang Huo Xiang	8.22	–
Gan Cao	8.24	–
She Gan	9.01	–
Dan Dou Chi	10.43	–
Gu Sui Bu	10.51	–
Zi Hua Di Ding	11.03	–
Ma Huang	11.49	–
Jin Yin Hua	12.22	–
Hong Jing Tian	12.42	–
Pi Pa Ye	12.53	–
Dan Shen	12.86	–
Qiang Huo	12.87	–
Yin Yang Huo	13.09	–
Shan Yao	13.44	–
Ce Bai Ye	13.75	–
Ye Jiao Teng	14.04	–
Tian Hua Fen	14.12	–
Chai Hu	14.43	–
Mu Zei	14.48	–
Ren Dong Teng	14.57	–
Jing Jie	15.68	–
Che Qian Zi	16.09	–
Lu Gen	16.20	–
Yu Jin	16.34	–
Zhi Ke	16.60	–
Bai Shao Yao	16.66	–
Ban Lan Gen	16.81	–
Zhi Gan Cao	17.76	–
Bai Xian Pi	17.87	–
Qing Pi	17.94	–
Zhu Ling	18.17	–
Chuan Xiong	18.40	–
(Fen) Bi Xie	18.71	–
Huang Lian	18.88	–
Suan Zao Ren	19.29	–
He Huan Pi	19.89	–
Ji Xue Teng	20.03	–
Ze Xie	20.08	–
Jiao Gu Lan	20.23	–
Tu Fu Ling	20.34	–
Sha Ren	20.35	–
Lian Qiao	20.44	–
Wu Wei Zi	20.62	–
Du Zhong	21.17	–
Ren Shen	21.21	–
Mao Dong Qing	21.25	–
Bo He	21.46	–
Shen Qu	21.64	–
Gua Lou	22.04	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ji Li	22.35	–
Zhe Bei Mu	22.48	–
Tao Ren	22.53	–
Huang Bai	22.95	–
Lian Zi	23.00	–
Cang Zhu	24.14	–
Ling Zhi	24.64	–
Gou Qi Zi	24.81	–
Sheng Jiang	25.12	–
Huo Ma Ren	25.91	–
Di Gu Pi	27.19	–
Yuan Zhi	27.24	–
Jie Geng	27.26	–
Hong Hua	27.33	–
Gou Teng	27.33	–
Tai Zi Shen	27.64	–
Gui Zhi	27.80	–
Huang Qin	27.91	–
Fu Ling	28.16	–
Qing Hao	28.79	–
Chen Pi	28.81	–
Rou Gui	29.67	–
Bai Jiang Cao	29.94	–
Sang Ye	30.37	–
Bai Zi Ren	30.55	–
Mang Xiao	31.09	–
Dang Gui	31.43	–
Chuang Mu Xiang	31.49	–
Ku Shen	31.51	–
Yu Zhu	32.15	–
Lai Fu Zi	32.66	–
Ban Xia (Jiang)	33.39	–
Chi Shao (Yao)	33.52	–
San Qi	33.69	–
Shan Yu Rou	33.78	–
Xie Bai	35.33	–
Ci Wu Jia	35.37	–
Ma Huang Gen	35.51	–
Dang Gui Wei	35.52	–
Pu Gong Ying	37.65	–
Zhu Ru	37.67	–
Cang Er Zi	38.81	–
Sang Zhi	39.78	–
Yi Yi Ren	39.95	–
Ze Lan	40.49	–
Fu Xiao Mai	42.78	–
Zhi Mu	43.32	–
Long Yan Rou	43.75	–
Fu Shen	44.80	–
Fo Shou	45.31	–
Ban Zhi Lian	45.45	–
Xi Xian Cao	46.66	–
Mu Gua	47.02	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fu Pen Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62855	62855	0.00	6.75
62856	62856	0.00	5.93

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fu Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60004-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fu Shen; Poriae cocos sclerotium pararadicis

Special notes

When selecting the *Fu Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fu Shen	1	0	2

Second-stage model

For differentiation of the substance/substance group *Fu Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fu Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fu Shen	G201HS212SH1	62617	40	from supplier
PhytoComm	Fu Shen	G201HS212SH1	62618	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Fu Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Fu Shen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fu Shen	G201HS212SH1	62617†	20
PhytoComm	Fu Shen	G201HS212SH1	62618†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Fu Shen*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Fu Shen	g201h1021321	1
Phytocomm	Fu Shen	G201H1021321	1
Phytocomm	Fu Shen	G201H1021422	2

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fu Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fu Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	1	40	0	12 436
Type C	1	0	4	852

The substance/substance group *Fu Shen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9952 % (> 99.9655 %)	100.0000 % (> 85.0000 %)
Type C	99.8837 % (> 99.2972 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fu Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Rou Gui	6.61	–
Yi Yi Ren	6.83	–
Ma Huang Gen	7.63	–
Fu Ling	8.97	–
Gui Zhi	13.31	–
Fu Xiao Mai	13.70	–
Sheng Jiang	14.01	–
He Huan Pi	16.31	–
Ji Xue Teng	17.69	–
Ban Xia (Jiang)	18.86	–
Gou Teng	18.97	–
Zhu Ru	19.94	–
Bai Shao Yao	20.26	–
Mu Zei	21.61	–
Di Gu Pi	23.16	–
Bai Zi Ren	25.96	–
Bai Xian Pi	26.08	–
Fo Shou	26.93	–
Lian Zi	26.96	–
Ji Li	27.74	–
Ling Zhi	28.11	–
Tao Ren	29.73	–
Gu Sui Bu	29.96	–
Ye Jiao Teng	30.78	–
Huo Ma Ren	31.74	–
Ci Wu Jia	31.92	–
Shen Qu	31.95	–
Lai Fu Zi	32.06	–
Mang Xiao	32.66	–
Lu Gen	32.99	–
Tai Zi Shen	33.01	–
Gua Lou	36.36	–
Zhu Ling	36.57	–
Tian Hua Fen	36.84	–
Ren Dong Teng	37.36	–
Zhe Bei Mu	38.16	–
Yuan Zhi	38.57	–
Chen Pi	38.87	–
Ze Xie	39.01	–
Tu Fu Ling	39.03	–
Zhi Mu	40.83	–
Chuan Xiong	41.00	–
Pi Pa Ye	41.23	–
Dan Dou Chi	41.41	–
Yu Jin	41.50	–
Suan Zao Ren	42.20	–
Ren Shen	42.56	–
Shan Yao	43.21	–
She Gan	43.82	–
(Fen) Bi Xie	44.30	–
Lian Qiao	44.37	–
Shan Yu Rou	45.29	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yan Hu Suo	45.34	–
Dan Shen	45.49	–
Jiao Gu Lan	47.04	–
Zhi Ke	48.08	–
Dang Gui	48.11	–
Jie Geng	49.52	–
Huang Qin	49.54	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fu Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62617	62617	0.00	6.61
62618	62618	0.00	6.64

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fu Xiao Mai**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60081-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fu Xiao Mai; Tritici aestivi semen levis

Special notes

When selecting the *Fu Xiao Mai* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fu Xiao Mai	3	0	0

Second-stage model

For differentiation of the substance/substance group *Fu Xiao Mai* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fu Xiao Mai*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fu Xiao Mai	G244HS021RT1	62423	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021RT1	62424	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021SK1	62631	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021SK1	62632	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021SW1	62801	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021SW1	62802	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Fu Xiao Mai*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Fu Xiao Mai*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fu Xiao Mai	G244HS021RT1	62423†	20
PhytoComm	Fu Xiao Mai	G244HS021RT1	62424†	20
PhytoComm	Fu Xiao Mai	G244HS021SK1	62631†	20
PhytoComm	Fu Xiao Mai	G244HS021SK1	62632†	20
PhytoComm	Fu Xiao Mai	G244HS021SW1	62801†	20
PhytoComm	Fu Xiao Mai	G244HS021SW1	62802†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Fu Xiao Mai*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fu Xiao Mai* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fu Xiao Mai* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	8	229	11	24 352
Type B	19	103	17	12 338
Type C	0	0	0	857

The substance/substance group *Fu Xiao Mai* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9798 % (> 99.9647 %)	95.4167 % (> 94.1667 %)
Type B	99.8714 % (> 99.8415 %)	85.8333 % (> 83.3333 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fu Xiao Mai* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yi Yi Ren	2.60	–
Zhu Ru	3.43	–
Sheng Jiang	3.75	–
Rou Gui	5.62	–
Di Gu Pi	7.24	–
Gui Zhi	7.26	–
Fu Ling	7.55	–
Bai Shao Yao	8.73	–
Ci Wu Jia	9.42	–
He Huan Pi	9.61	–
Ban Xia (Jiang)	10.27	–
Ji Li	11.10	–
Lian Zi	11.78	–
Fu Shen	12.00	–
Shen Qu	12.48	–
Bai Zi Ren	13.31	–
Tao Ren	13.82	–
Ma Huang Gen	14.35	–
Ling Zhi	14.79	–
Lai Fu Zi	14.92	–
Gou Teng	16.29	–
Ji Xue Teng	16.54	–
Gua Lou	16.64	–
Mu Zei	16.76	–
Zhe Bei Mu	17.02	–
Bai Xian Pi	18.19	–
Yuan Zhi	18.63	–
Ye Jiao Teng	19.29	–
Shan Yao	19.96	–
Huo Ma Ren	20.02	–
Fo Shou	20.03	–
Tian Hua Fen	20.14	–
Tai Zi Shen	21.30	–
Gu Sui Bu	21.92	–
Pi Pa Ye	22.49	–
Yu Jin	22.85	–
Tu Fu Ling	25.57	–
Ren Dong Teng	26.23	–
Lu Gen	26.68	–
Ze Xie	26.94	–
Lian Qiao	28.52	–
Suan Zao Ren	28.95	–
Chuan Xiong	29.77	–
Zhu Ling	30.28	–
Shan Yu Rou	31.14	–
Zhi Mu	31.27	–
Chen Pi	32.42	–
Yan Hu Suo	32.87	–
San Qi	33.51	–
(Fen) Bi Xie	33.73	–
Ren Shen	33.76	–
Mang Xiao	34.08	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Cang Zhu	34.32	–
Jin Yin Hua	34.38	–
Fu Zi	34.57	–
Ce Bai Ye	34.74	–
Jie Geng	35.10	–
Dang Gui Wei	35.66	–
Jiao Gu Lan	36.38	–
Zhi Ke	37.09	–
Fu Pen Zi	37.14	–
Mao Dong Qing	37.88	–
Dan Dou Chi	38.46	–
Yu Zhu	38.87	–
Dan Shen	39.13	–
Chai Hu	39.13	–
Dang Gui	39.48	–
She Gan	39.89	–
Huang Qin	40.01	–
Che Qian Zi	41.84	–
Wu Wei Zi	41.94	–
Yin Yang Huo	42.44	–
Hou Po	42.90	–
Mu Gua	43.66	–
Ku Shen	44.61	–
Ban Lan Gen	44.77	–
Guang Huo Xiang	47.60	–
Gou Qi Zi	48.76	–
Mai Men Dong	48.84	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fu Xiao Mai* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62423	62423	0.00	4.57
62424	62424	0.00	2.60
62631	62631	0.00	4.41
62632	62632	0.00	2.80
62801	62801	0.00	4.16
62802	62802	0.00	3.75

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances,

thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Fu Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50884-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Fu Zi; Aconiti radix lateralis praep.

Special notes

When selecting the *Fu Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Fu Zi	2	0	3

Second-stage model

For differentiation of the substance/substance group *Fu Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Fu Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fu Zi	G005HS158SP1	62699	40	from supplier
PhytoComm	Fu Zi	G005HS158SP1	62700	40	from supplier
PhytoComm	Fu Zi	G005HS158TH1	62943	40	from supplier
PhytoComm	Fu Zi	G005HS158TH1	62944	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Fu Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Fu Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Fu Zi	G005HS158SP1	62699†	20
PhytoComm	Fu Zi	G005HS158SP1	62700†	20
PhytoComm	Fu Zi	G005HS158TH1	62943†	20
PhytoComm	Fu Zi	G005HS158TH1	62944†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 4 batches from the substance/substance group *Fu Zi*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Fu Zi	g005h0801122	1
phytocomm	Fu Zi	g005h0801322	1
Phytocomm	Fu Zi	G005H0801322	1
Phytocomm	Fu Zi	G005H0801524	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Fu Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Fu Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	156	4	24 440
Type B	0	73	7	12 397
Type C	0	0	4	853

The substance/substance group *Fu Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	97.5000 % (> 95.6250 %)
Type B	100.0000 % (> 99.9403 %)	91.2500 % (> 87.5000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Fu Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shan Yao	3.39	–
Pi Pa Ye	4.56	–
Zi Hua Di Ding	4.89	–
Fu Pen Zi	5.05	–
Zhe Bei Mu	5.98	–
Tian Hua Fen	6.21	–
Ren Dong Teng	7.03	–
Lian Qiao	7.33	–
Ze Xie	7.46	–
Ji Li	8.09	–
Dan Shen	8.32	–
Gan Cao	8.78	–
Hou Po	8.83	–
Gu Sui Bu	8.83	–
Mu Zei	8.83	–
Gua Lou	9.79	–
He Huan Pi	9.92	–
Jin Yin Hua	10.01	–
Ling Zhi	10.23	–
Guang Huo Xiang	10.43	–
Yan Hu Suo	10.46	–
Qiang Huo	10.46	–
Hong Jing Tian	10.47	–
Jing Jie	10.59	–
Bai Zi Ren	10.66	–
Chuan Xiong	10.68	–
Shen Qu	10.70	–
Huo Ma Ren	11.04	–
Yin Yang Huo	11.94	–
Dan Dou Chi	11.97	–
Ban Lan Gen	12.34	–
Bai Shao Yao	12.69	–
Bo He	12.87	–
Yu Jin	13.12	–
Chai Hu	13.45	–
Che Qian Zi	13.57	–
She Gan	13.60	–
Ye Jiao Teng	14.01	–
Ce Bai Ye	14.04	–
Tai Zi Shen	14.12	–
Ma Huang	14.17	–
Sheng Jiang	14.24	–
Lian Zi	14.30	–
Ji Xue Teng	14.34	–
Lai Fu Zi	14.55	–
Jiao Gu Lan	14.79	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Ren	14.79	–
Huang Bai	14.93	–
Cang Zhu	15.41	–
Fu Ling	15.69	–
Gui Zhi	15.77	–
Suan Zao Ren	16.01	–
Chen Pi	16.03	–
Bai Xian Pi	16.03	–
Zhu Ling	16.42	–
Yuan Zhi	16.45	–
Huang Lian	16.52	–
Mao Dong Qing	16.57	–
Tu Fu Ling	16.66	–
Zhi Ke	16.68	–
Zhi Gan Cao	16.75	–
Lu Gen	17.04	–
Du Zhong	17.29	–
Ren Shen	17.42	–
(Fen) Bi Xie	17.83	–
Tao Ren	18.51	–
Qing Pi	18.77	–
Gou Teng	18.84	–
San Qi	19.84	–
Wu Wei Zi	19.95	–
Rou Gui	20.13	–
Zhu Ru	20.26	–
Di Gu Pi	20.26	–
Dang Gui	20.48	–
Ku Shen	20.58	–
Yu Zhu	21.15	–
Jie Geng	22.50	–
Ci Wu Jia	22.63	–
Qing Hao	22.91	–
Hong Hua	23.06	–
Ban Xia (Jiang)	23.14	–
Gou Qi Zi	23.29	–
Bai Jiang Cao	23.71	–
Dang Gui Wei	23.84	–
Fu Xiao Mai	24.84	–
Xie Bai	26.63	–
Yi Yi Ren	26.70	–
Shan Yu Rou	27.51	–
Ma Huang Gen	27.55	–
Huang Qin	27.59	–
Zhi Mu	27.76	–
Fo Shou	29.07	–
Chuang Mu Xiang	29.60	–
Mang Xiao	30.58	–
Pu Gong Ying	30.58	–
Sang Ye	30.62	–
Chi Shao (Yao)	31.74	–
Long Yan Rou	33.46	–
Mu Gua	33.65	–
Sang Zhi	35.05	–
Ze Lan	35.12	–
Cang Er Zi	37.28	–
Ban Zhi Lian	39.82	–
Fu Shen	40.52	–
Xi Xian Cao	43.14	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	48.19	–
Jiang Huang	49.16	–
(Shi) Chang Pu	49.23	–
Nü Zhen Zi	49.89	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fu Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62699	62699	0.00	3.66
62700	62700	0.00	3.39
62943	62943	0.00	5.05
62944	62944	0.00	4.89

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gan Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60011-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gan Cao; Glycyrrhizae radix

Special notes

When selecting the *Gan Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gan Cao	3	0	3

Second-stage model

For differentiation of the substance/substance group *Gan Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gan Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gan Cao	G119H0521822	62341	40	from supplier
PhytoComm	Gan Cao	G119H0521822	62342	40	from supplier
PhytoComm	Gan Cao	G119H0521823	62533	40	from supplier
PhytoComm	Gan Cao	G119H0521823	62534	40	from supplier
PhytoComm	Gan Cao	G119HS089TL1	62879	40	from supplier
PhytoComm	Gan Cao	G119HS089TL1	62880	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Gan Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Gan Cao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gan Cao	G119H0521822	62341†	20
PhytoComm	Gan Cao	G119H0521822	62342†	20
PhytoComm	Gan Cao	G119H0521823	62533†	20
PhytoComm	Gan Cao	G119H0521823	62534†	20
PhytoComm	Gan Cao	G119HS089TL1	62879†	20
PhytoComm	Gan Cao	G119HS089TL1	62880†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 18 spectra from 7 *Apo-Ident* customers from 6 batches from the substance/substance group *Gan Cao*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gan Cao	g119h0521121	1
Phytocomm	Gan Cao	G119H0521121	1
PhytoComm	Gan Cao	G119H0521121	4
Phytocomm	Gan Cao	G119H0521522	8
PhytoComm	Gan Cao	G119H0521524	1
PhytoComm	Gan Cao	G119H0521822	2
PhytoComm	Gan Cao	G119H0521823	1

- 839 spectra from 13 *Apo-Ident* customers from a total of 513 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gan Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gan Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	12	6	839

The substance/substance group *Gan Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8244 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gan Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	6.29	–
Zhi Gan Cao	7.73	–
Hong Jing Tian	8.84	–
Fu Zi	9.11	–
Jin Yin Hua	9.16	–
Fu Pen Zi	9.45	–
Dan Shen	9.55	–
Hou Po	9.66	–
Jiao Gu Lan	10.85	–
Zhi Ke	10.93	–
Dan Dou Chi	11.85	–
Ma Huang	12.30	–
Ban Lan Gen	12.87	–
Qiang Huo	12.88	–
Yin Yang Huo	13.61	–
Yan Hu Suo	14.50	–
Qing Pi	14.68	–
Pi Pa Ye	14.90	–
Chai Hu	14.91	–
Ye Jiao Teng	15.24	–
Bo He	15.32	–
Huang Bai	15.61	–
Shan Yao	16.03	–
Chuan Xiong	16.37	–
Ren Dong Teng	16.82	–
Bai Shao Yao	16.90	–
Tian Hua Fen	17.17	–
She Gan	17.29	–
Guang Huo Xiang	17.58	–
Gu Sui Bu	17.78	–
Suan Zao Ren	18.67	–
Qing Hao	18.79	–
Hong Hua	18.81	–
Jing Jie	18.87	–
Mao Dong Qing	19.00	–
Sha Ren	19.16	–
Gua Lou	19.29	–
Chi Shao (Yao)	19.35	–
Ce Bai Ye	19.66	–
Mu Zei	19.72	–
Bai Xian Pi	19.76	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lu Gen	19.91	–
Yu Jin	19.92	–
Wu Wei Zi	20.47	–
Chuang Mu Xiang	20.61	–
Du Zhong	20.67	–
Gou Qi Zi	20.81	–
Huang Lian	20.97	–
Shan Yu Rou	21.81	–
Lian Zi	22.10	–
Lian Qiao	22.17	–
Ji Xue Teng	22.30	–
Cang Zhu	22.31	–
Che Qian Zi	22.49	–
Ji Li	22.92	–
Ren Shen	23.30	–
Yuan Zhi	24.00	–
He Huan Pi	24.19	–
Shen Qu	24.71	–
(Fen) Bi Xie	25.83	–
Ling Zhi	26.01	–
Bai Jiang Cao	26.32	–
Tu Fu Ling	26.63	–
Zhe Bei Mu	26.65	–
Xiang Fu	26.67	–
Di Gu Pi	26.74	–
Ze Xie	26.88	–
Zhu Ling	27.26	–
Huang Qin	27.76	–
Ku Shen	28.40	–
Mang Xiao	28.55	–
Niu Bang Zi	28.63	–
Tao Ren	28.65	–
Chen Pi	29.07	–
Sang Ye	29.16	–
Xie Bai	29.35	–
Dang Gui	30.60	–
Sheng Jiang	30.88	–
Jie Geng	31.23	–
Gou Teng	31.38	–
Yu Zhu	32.26	–
Fu Ling	32.57	–
Rou Gui	32.92	–
Cang Er Zi	33.11	–
Tai Zi Shen	33.19	–
Bai Zi Ren	33.29	–
Huo Ma Ren	33.58	–
Gui Zhi	34.08	–
Pu Gong Ying	35.29	–
Dang Gui Wei	35.30	–
San Qi	35.53	–
Zi Su Zi	35.77	–
Chuan Lian Zi	36.30	–
Mu Gua	36.68	–
Ma Huang Gen	38.20	–
Long Yan Rou	38.50	–
Ze Lan	39.26	–
(Huai) Niu Xi	39.52	–
Sang Zhi	40.10	–
Ci Wu Jia	40.12	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lai Fu Zi	41.26	–
Huang Qi	41.56	–
Ban Xia (Jiang)	42.26	–
Ban Zhi Lian	42.63	–
Zhi Mu	44.36	–
Yi Yi Ren	44.94	–
Xi Xian Cao	45.78	–
Zhu Ru	46.19	–
Wu Mei	47.11	–
Chuan Mu Tong	49.01	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gan Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62341	62341	0.00	19.16
62342	62342	0.00	18.81
62533	62533	0.00	21.76
62534	62534	0.00	22.09
62879	62879	0.00	6.29
62880	62880	0.00	7.26

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gan Jiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10004524-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gan Jiang; Zingiberis rhizoma

Special notes

When selecting the *Gan Jiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gan Jiang	1	0	4

Second-stage model

For differentiation of the substance/substance group *Gan Jiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gan Jiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gan Jiang	G252H1145021	62805	40	from supplier
PhytoComm	Gan Jiang	G252H1145021	62806	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Gan Jiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Gan Jiang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gan Jiang	G252H1145021	62805†	20
PhytoComm	Gan Jiang	G252H1145021	62806†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Gan Jiang*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gan Jiang	g252h01145221	1
Phytocomm	Gan Jiang	G252H1145221	1
Phytocomm	Gan Jiang	G252H1145421	1
Phytocomm	Gan Jiang	G252H1145521	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gan Jiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gan Jiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	3	0	4	850

The substance/substance group *Gan Jiang* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.6512 % (> 99.0646 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gan Jiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zi Su Zi	6.46	–
Sha Shen (Bei)	11.86	–
Tian Hua Fen	13.95	–
Cang Er Zi	14.54	–
Chuan Lian Zi	14.99	–
Di Gu Pi	16.15	–
Niu Bang Zi	16.32	–
Ji Li	16.95	–
Yan Hu Suo	17.02	–
Sang Zhi	17.27	–
(Shi) Chang Pu	18.07	–
(Huai) Niu Xi	18.75	–
Gua Lou	20.13	–
Jie Geng	20.59	–
Mi Huan Jun	22.38	–
E Zhu	23.71	–
Jiang Huang	24.28	–
Shan Yao	24.76	–
Dang Gui	25.06	–
Suan Zao Ren	26.26	–
Lai Fu Zi	26.72	–
(Bai) Dou Kou	26.90	–
Mang Xiao	27.85	–
Xiao Hui Xiang	28.02	–
Chuan Mu Tong	28.07	–
Sha Ren	28.42	–
Ku Shen	28.61	–
Lian Qiao	28.62	–
Zhi Gan Cao	29.24	–
Huang Qi	29.48	–
Xiang Fu	29.85	–
Chuan Niu Xi	30.17	–
Chen Pi	30.38	–
Bai Zhi	31.07	–
Long Dan (Cao)	31.32	–
Mai Ya	32.74	–
Bai Zhu	34.74	–
Pu Gong Ying	35.09	–
Bai He	35.30	–
Qin Jiao	36.20	–
Shan Yu Rou	38.31	–
Yi Mu Cao	38.93	–
Bai Hua She She Cao	41.80	–
Wu Yao	42.63	–
Ban Zhi Lian	43.01	–
Wu Wei Zi	43.32	–
Qiang Huo	43.75	–
He Huan Pi	44.10	–
Da Zao	44.56	–
Tu Fu Ling	44.83	–
Sang Ji Shend	45.17	–
Dong Gua Zi	45.20	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	45.23	–
Jin Yin Hua	45.37	–
Huang Lian	45.42	–
Du Huo	45.69	–
Fang Feng	46.12	–
Fu Ling	46.27	–
Zi Hua Di Ding	46.27	–
Chi Shao (Yao)	46.57	–
Lian Zi	46.81	–
Bai Xian Pi	47.09	–
Mu Dan Pi	47.23	–
Bo He	47.57	–
Ba Ji Tian	48.46	–
Yuan Zhi	48.80	–
Ju Hua	49.27	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gan Jiang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62805	62805	0.00	6.46
62806	62806	0.00	6.67

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ge Gen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60050-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ge Gen; Puerariae radix

Special notes

When selecting the *Ge Gen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ge Gen	1	0	3

Second-stage model

For differentiation of the substance/substance group *Ge Gen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ge Gen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ge Gen	G206H1301822	62489	40	from supplier
PhytoComm	Ge Gen	G206H1301822	62490	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ge Gen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ge Gen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ge Gen	G206H1301822	62489†	20
PhytoComm	Ge Gen	G206H1301822	62490†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Ge Gen*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ge Gen	G206H1301221	2
PhytoComm	Ge Gen	G206H1301323	1
Phytocomm	Ge Gen	G206H1301422	2

- 852 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ge Gen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ge Gen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	5	852

The substance/substance group *Ge Gen* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ge Gen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	18.93	–
Sang Ji Shend	23.80	–
Gu Sui Bu	38.78	–
Jing Jie	38.90	–
He Shou Wu	40.45	–
Hong Jing Tian	40.47	–
Gou Teng	41.02	–
Du Zhong	47.20	–
Hua Shi	51.03	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ge Gen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62489	62489	0.00	19.01
62490	62490	0.00	18.93

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gou Qi Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60094-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gou Qi Zi; Lycii chinensis fructus

Special notes

When selecting the *Gou Qi Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gou Qi Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Gou Qi Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gou Qi Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gou Qi Zi	G150HS171SK1	62615	40	from supplier
PhytoComm	Gou Qi Zi	G150HS171SK1	62616	40	from supplier
PhytoComm	Gou Qi Zi	G150HS171TH1	62875	40	from supplier
PhytoComm	Gou Qi Zi	G150HS171TH1	62876	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Gou Qi Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Gou Qi Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gou Qi Zi	G150HS171SK1	62615†	20
PhytoComm	Gou Qi Zi	G150HS171SK1	62616†	20
PhytoComm	Gou Qi Zi	G150HS171TH1	62875†	20
PhytoComm	Gou Qi Zi	G150HS171TH1	62876†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Gou Qi Zi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gou Qi Zi	g150h0913221	1
PhytoComm	Gou Qi Zi	G150H0913221	1
Phytocomm	Gou Qi Zi	G150H0913422	4

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gou Qi Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gou Qi Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	6	851

The substance/substance group *Gou Qi Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gou Qi Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui	6.96	–
Cang Zhu	7.51	–
Chuang Mu Xiang	8.79	–
Chen Pi	9.35	–
Yu Zhu	11.81	–
Long Yan Rou	12.15	–
Jie Geng	12.26	–
Hong Jing Tian	12.61	–
Shan Yu Rou	12.88	–
Ban Lan Gen	13.69	–
Dang Gui Wei	13.72	–
Gua Lou	13.76	–
Mu Gua	13.81	–
Chuan Xiong	14.36	–
Ren Shen	14.82	–
Xie Bai	16.91	–
Zhi Ke	16.91	–
Ku Shen	16.92	–
Gan Cao	17.13	–
Ze Xie	17.28	–
Jin Yin Hua	17.57	–
Dan Dou Chi	17.83	–
Jiao Gu Lan	18.03	–
Tian Hua Fen	18.04	–
San Qi	18.48	–
Yuan Zhi	18.65	–
Ma Huang	18.70	–
Mu Zei	18.89	–
Lian Qiao	19.45	–
Ren Dong Teng	19.70	–
Zhi Mu	20.33	–
Pi Pa Ye	20.80	–
Tai Zi Shen	20.83	–
Suan Zao Ren	20.85	–
Bai Shao Yao	20.93	–
Fu Pen Zi	20.98	–
Shan Yao	21.45	–
Hou Po	21.46	–
Dan Shen	21.86	–
Zi Hua Di Ding	22.08	–
Zhe Bei Mu	22.30	–
Fu Zi	22.86	–
Ye Jiao Teng	22.98	–
Tao Ren	23.28	–
Lu Gen	23.34	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ling Zhi	23.39	–
He Huan Pi	23.44	–
Huang Qin	23.79	–
Hong Hua	24.41	–
Di Gu Pi	24.65	–
Guang Huo Xiang	25.17	–
Yan Hu Suo	25.18	–
Wu Wei Zi	25.29	–
Yin Yang Huo	25.30	–
Gu Sui Bu	25.71	–
Lian Zi	25.83	–
Shen Qu	26.33	–
Lai Fu Zi	27.00	–
Gou Teng	27.37	–
Gui Zhi	27.44	–
Zhi Gan Cao	27.49	–
Ji Li	27.50	–
Bai Zi Ren	28.20	–
Ci Wu Jia	28.49	–
Fo Shou	28.91	–
Chai Hu	29.63	–
Ji Xue Teng	29.98	–
Mang Xiao	30.06	–
She Gan	30.71	–
Huo Ma Ren	31.62	–
Tu Fu Ling	31.66	–
Ban Xia (Jiang)	32.24	–
Che Qian Zi	32.33	–
Qiang Huo	32.50	–
Bai Xian Pi	32.60	–
Sheng Jiang	32.87	–
Chi Shao (Yao)	32.87	–
Sang Zhi	32.92	–
Rou Gui	33.09	–
Zhu Ru	33.30	–
Mao Dong Qing	33.50	–
Bo He	34.42	–
Fu Ling	35.15	–
Qing Pi	36.32	–
Mai Men Dong	36.51	–
(Fen) Bi Xie	36.55	–
Ce Bai Ye	36.61	–
Jing Jie	37.59	–
Sha Ren	38.20	–
Yu Jin	38.27	–
Huang Lian	38.50	–
Huang Bai	38.67	–
Huang Qi	40.93	–
Cang Er Zi	40.96	–
Fu Xiao Mai	41.14	–
Zhu Ling	41.89	–
Chuan Lian Zi	42.23	–
Yi Yi Ren	42.56	–
Ma Huang Gen	42.69	–
Sang Ye	45.11	–
Du Zhong	45.13	–
Niu Bang Zi	46.00	–
Fu Shen	48.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gou Qi Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62615	62615	0.00	7.22
62616	62616	0.00	6.96
62875	62875	0.00	9.56
62876	62876	0.00	8.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gou Teng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60097-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gou Teng; Uncariae ramulus cum uncis

Special notes

When selecting the *Gou Teng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gou Teng	2	0	2

Second-stage model

For differentiation of the substance/substance group *Gou Teng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gou Teng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gou Teng	G247H1210822	62313	40	from supplier
PhytoComm	Gou Teng	G247H1210822	62314	40	from supplier
PhytoComm	Gou Teng	G247HS288SG1	62629	40	from supplier
PhytoComm	Gou Teng	G247HS288SG1	62630	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Gou Teng*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Gou Teng*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gou Teng	G247H1210822	62313†	20
PhytoComm	Gou Teng	G247H1210822	62314†	20
PhytoComm	Gou Teng	G247HS288SG1	62629†	20
PhytoComm	Gou Teng	G247HS288SG1	62630†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Gou Teng*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gou Teng	g247h1210122	1
Phytocomm	Gou Teng	G247H1210422	2

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gou Teng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gou Teng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	6	159	1	24 434
Type B	0	77	3	12 397
Type C	2	0	3	852

The substance/substance group *Gou Teng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9893 % (> 99.9743 %)	99.3750 % (> 97.5000 %)
Type B	100.0000 % (> 99.9403 %)	96.2500 % (> 92.5000 %)
Type C	99.5349 % (> 98.9489 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gou Teng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Rou Gui	3.07	–
He Huan Pi	4.58	–
Fu Ling	5.73	–
Gui Zhi	5.81	–
Ma Huang Gen	6.54	–
Yi Yi Ren	6.56	–
Mu Zei	6.88	–
Ji Xue Teng	7.70	–
Huo Ma Ren	8.05	–
Sheng Jiang	8.49	–
Lu Gen	9.41	–
Fu Shen	10.13	–
Fu Xiao Mai	10.28	–
Ci Wu Jia	11.08	–
Tai Zi Shen	11.81	–
Tao Ren	12.10	–
Ban Xia (Jiang)	12.53	–
Ren Dong Teng	13.23	–
Gu Sui Bu	13.32	–
Bai Shao Yao	13.36	–
Zhu Ling	13.37	–
Bai Xian Pi	13.59	–
Lai Fu Zi	13.93	–
Ling Zhi	14.37	–
Lian Zi	14.45	–
(Fen) Bi Xie	15.27	–
Fo Shou	15.30	–
Ye Jiao Teng	15.51	–
Di Gu Pi	16.06	–
She Gan	16.42	–
Ji Li	17.20	–
Zhe Bei Mu	17.38	–
Bai Zi Ren	18.01	–
Shen Qu	18.24	–
Zhu Ru	18.38	–
Suan Zao Ren	18.72	–
Nü Zhen Zi	21.08	–
Tu Fu Ling	21.15	–
Dan Dou Chi	21.23	–
Yu Jin	21.67	–
Tian Hua Fen	21.76	–
Yi Mu Cao	22.82	–
Mang Xiao	23.14	–
Chuan Xiong	23.69	–
Yan Hu Suo	24.43	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Shen	24.77	–
Lian Qiao	25.12	–
Ce Bai Ye	25.38	–
Gua Lou	25.52	–
Ze Xie	25.90	–
Jing Jie	26.39	–
Chen Pi	27.30	–
Pi Pa Ye	27.53	–
Yin Yang Huo	28.43	–
Shan Yao	28.73	–
Mao Dong Qing	28.92	–
Yuan Zhi	29.10	–
Chai Hu	29.43	–
Xuan Fu Hua	30.74	–
Jiao Gu Lan	31.01	–
Hong Jing Tian	31.34	–
Guang Huo Xiang	31.70	–
Jin Yin Hua	31.74	–
Fu Zi	33.53	–
Zhi Ke	33.84	–
Fu Pen Zi	34.02	–
Yu Zhu	34.10	–
Ma Huang	34.12	–
Ban Lan Gen	35.28	–
Zhi Shi	35.79	–
Sang Ji Shend	36.01	–
Gan Cao	36.45	–
Ren Shen	36.57	–
Cang Zhu	36.64	–
Jie Geng	37.31	–
Hou Po	37.38	–
Qiang Huo	37.57	–
Che Qian Zi	38.14	–
Dang Gui Wei	38.57	–
Dang Gui	39.06	–
Wu Zhu Yu	39.89	–
Huang Qin	40.03	–
Xin Yi	40.67	–
Zhi Mu	41.07	–
Zi Hua Di Ding	41.18	–
Sha Ren	42.30	–
Huang Lian	42.53	–
Shan Yu Rou	42.55	–
San Qi	42.62	–
Wu Wei Zi	42.84	–
Xuan Shen	43.75	–
Jin Qian Cao	43.75	–
Sang Ye	44.11	–
Wu Jia Pi	44.14	–
Ding Xiang	44.16	–
Sang Zhi	44.57	–
Sang Bai Pi	45.72	–
Rou Cong Rong	46.45	–
Long Yan Rou	46.46	–
Xia Ku Cao	46.83	–
Mu Gua	47.62	–
Dan Zhu Ye	48.49	–
Du Zhong	48.71	–
Chuang Mu Xiang	48.83	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Hua She She Cao	49.22	–
Gou Qi Zi	49.45	–
Han Lian Cao	49.69	–
Hu Zhang	49.70	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gou Teng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62313	62313	0.00	21.08
62314	62314	0.00	21.99
62629	62629	0.00	5.73
62630	62630	0.00	3.07

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gu Sui Bu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002263-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gu Sui Bu; Drynariae rhizoma

Special notes

When selecting the *Gu Sui Bu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gu Sui Bu	2	0	3

Second-stage model

For differentiation of the substance/substance group *Gu Sui Bu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gu Sui Bu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gu Sui Bu	G094H1011921	62779	40	from supplier
PhytoComm	Gu Sui Bu	G094H1011921	62780	40	from supplier
PhytoComm	Gu Sui Bu	G094HS230TK1	62921	40	from supplier
PhytoComm	Gu Sui Bu	G094HS230TK1	62922	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Gu Sui Bu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Gu Sui Bu*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gu Sui Bu	G094H1011921	62779†	20
PhytoComm	Gu Sui Bu	G094H1011921	62780†	20
PhytoComm	Gu Sui Bu	G094HS230TK1	62921†	20
PhytoComm	Gu Sui Bu	G094HS230TK1	62922†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 5 *Apo-Ident* customers from 4 batches from the substance/substance group *Gu Sui Bu*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gu Sui Bu	g094h1011023	1
Phytocomm	Gu Sui Bu	g094h1011221	1
Phytocomm	Gu Sui Bu	G094H1011221	2
PhytoComm	Gu Sui Bu	G094H1011221	1
PhytoComm	Gu Sui Bu	G094H1011522	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gu Sui Bu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gu Sui Bu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	3	0	6	848

The substance/substance group *Gu Sui Bu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.4574 % (> 98.8703 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gu Sui Bu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shen Qu	5.24	–
Ren Dong Teng	7.08	–
Mu Zei	7.41	–
Huo Ma Ren	7.72	–
Fu Zi	8.55	–
She Gan	9.05	–
Ce Bai Ye	9.43	–
Pi Pa Ye	9.69	–
Chuan Xiong	9.87	–
(Fen) Bi Xie	10.21	–
Lu Gen	10.32	–
Fu Pen Zi	10.42	–
He Huan Pi	10.42	–
Shan Yao	10.47	–
Tian Hua Fen	11.22	–
Bai Shao Yao	11.24	–
Ji Li	11.27	–
Zhu Ling	11.30	–
Gui Zhi	11.43	–
Ling Zhi	11.53	–
Ji Xue Teng	11.60	–
Bai Xian Pi	11.70	–
Ye Jiao Teng	11.73	–
Dan Dou Chi	11.78	–
Hou Po	11.78	–
Yan Hu Suo	11.81	–
Gou Teng	12.12	–
Lian Qiao	12.28	–
Dan Shen	12.57	–
Fu Ling	12.60	–
Gua Lou	12.69	–
Lian Zi	12.83	–
Yu Jin	13.23	–
Ze Xie	13.36	–
Suan Zao Ren	13.69	–
Yin Yang Huo	13.90	–
Tao Ren	14.05	–
Zhe Bei Mu	14.13	–
Jin Yin Hua	14.22	–
Sang Ji Shend	14.59	–
Sheng Jiang	14.60	–
Che Qian Zi	14.72	–
Guang Huo Xiang	14.80	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Lan Gen	14.92	–
Tai Zi Shen	15.79	–
Chai Hu	15.88	–
Ma Huang	15.91	–
He Shou Wu	16.23	–
Sang Bai Pi	16.43	–
Bai Zi Ren	16.62	–
Ma Huang Gen	17.14	–
Ci Wu Jia	17.32	–
Hong Jing Tian	17.54	–
Gan Cao	17.75	–
Mao Dong Qing	17.78	–
Chen Pi	17.91	–
Tu Fu Ling	18.17	–
Rou Gui	18.23	–
Lai Fu Zi	18.55	–
Jiao Gu Lan	18.78	–
Qiang Huo	19.03	–
Fo Shou	19.62	–
Ban Xia (Jiang)	19.69	–
Yin Chen Hao	19.71	–
Di Gu Pi	20.14	–
Yuan Zhi	20.27	–
Cang Zhu	20.70	–
Yi Yi Ren	20.91	–
Dan Zhu Ye	21.34	–
Zhu Ru	22.03	–
Yu Zhu	22.69	–
Ren Shen	22.93	–
Fu Xiao Mai	23.14	–
Jing Jie	23.44	–
Jiang Huang	23.57	–
Zhi Gan Cao	24.16	–
Xuan Fu Hua	24.29	–
Dang Gui Wei	24.94	–
Zi Hua Di Ding	25.20	–
Wu Wei Zi	25.39	–
Jie Geng	25.48	–
Dang Gui	25.51	–
Mang Xiao	25.94	–
Wu Yao	26.75	–
Bu Gu Zhi	26.88	–
Zhi Ke	26.98	–
Qing Pi	27.75	–
Sha Ren	27.83	–
Gou Qi Zi	27.85	–
San Qi	28.29	–
Huang Lian	28.83	–
Huang Qin	29.28	–
Bo He	30.50	–
Tu Si Zi	30.60	–
Du Zhong	30.90	–
(Bai) Dou Kou	30.92	–
Wu Zhu Yu	31.42	–
Yi Mu Cao	31.78	–
Fu Shen	32.06	–
Mu Dan Pi	32.53	–
Hu Zhang	32.67	–
Long Yan Rou	33.05	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Mu	33.15	–
Ku Shen	33.23	–
Sang Zhi	33.80	–
(Shi) Chang Pu	33.96	–
Huang Bai	34.26	–
Chuang Mu Xiang	34.50	–
Xie Bai	34.60	–
Chi Shao (Yao)	35.36	–
Shan Yu Rou	35.44	–
Ge Gen	35.51	–
Nü Zhen Zi	36.51	–
E Zhu	38.29	–
Mu Gua	39.05	–
Hong Hua	40.32	–
Xia Ku Cao	40.86	–
Fang Feng	42.25	–
Ju Hua	42.77	–
Xiao Hui Xiang	43.13	–
Bai Jiang Cao	44.68	–
Xi Xian Cao	45.41	–
Cang Er Zi	45.96	–
Xin Yi	47.47	–
Bai Hua She She Cao	47.90	–
Qing Hao	47.90	–
Zhi Shi	48.46	–
Chuan Niu Xi	48.74	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gu Sui Bu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62779	62779	0.00	11.67
62780	62780	0.00	11.28
62921	62921	0.00	5.29
62922	62922	0.00	5.24

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gua Lou**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60128-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gua Lou; Trichosanthis fructus

Special notes

When selecting the *Gua Lou* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gua Lou	3	0	2

Second-stage model

For differentiation of the substance/substance group *Gua Lou* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gua Lou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gua Lou	G241H1096821	62317	40	from supplier
PhytoComm	Gua Lou	G241H1096821	62318	40	from supplier
PhytoComm	Gua Lou	G241HS202SK1	62547	40	from supplier
PhytoComm	Gua Lou	G241HS202SK1	62548	40	from supplier
PhytoComm	Gua Lou	G241HS202SV1	62797	40	from supplier
PhytoComm	Gua Lou	G241HS202SV1	62798	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Gua Lou*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Gua Lou*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gua Lou	G241H1096821	62317†	20
PhytoComm	Gua Lou	G241H1096821	62318†	20
PhytoComm	Gua Lou	G241HS202SK1	62547†	20
PhytoComm	Gua Lou	G241HS202SK1	62548†	20
PhytoComm	Gua Lou	G241HS202SV1	62797†	20
PhytoComm	Gua Lou	G241HS202SV1	62798†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Gua Lou*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gua Lou	g241h1096321	1
Phytocomm	Gua Lou	G241H1096321	1
Phytocomm	Gua Lou	G241H1096521	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gua Lou* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gua Lou* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	239	1	24 359
Type B	2	118	2	12 355
Type C	0	0	3	854

The substance/substance group *Gua Lou* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	99.5833 % (> 98.3333 %)
Type B	99.9714 % (> 99.9415 %)	98.3333 % (> 95.8333 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gua Lou* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui Wei	4.29	–
Fo Shou	4.37	–
Yuan Zhi	6.71	–
Bai Shao Yao	6.91	–
Jie Geng	7.74	–
Ye Jiao Teng	8.72	–
Zhi Mu	9.33	–
Tai Zi Shen	9.52	–
Shen Qu	9.57	–
Mu Zei	10.31	–
Ze Xie	11.07	–
He Huan Pi	11.20	–
Pi Pa Ye	11.21	–
Zhe Bei Mu	11.54	–
Ji Li	11.57	–
Yu Zhu	11.63	–
Di Gu Pi	11.67	–
Tao Ren	11.78	–
Ci Wu Jia	12.24	–
San Qi	12.34	–
Chen Pi	12.41	–
Tian Hua Fen	12.64	–
Bai Zi Ren	13.31	–
Cang Zhu	13.57	–
Jin Yin Hua	13.98	–
Rou Gui	14.06	–
Xiao Hui Xiang	14.11	–
Lai Fu Zi	14.21	–
Lian Zi	14.25	–
Chuan Xiong	14.41	–
Ban Xia (Jiang)	14.43	–
Dang Gui	14.62	–
Ling Zhi	14.80	–
Zhu Ru	14.83	–
Shan Yao	15.03	–
Ban Lan Gen	15.38	–
Ku Shen	15.58	–
Mu Gua	15.68	–
Niu Bang Zi	15.70	–
Zi Su Zi	15.84	–
Sheng Jiang	15.85	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Dong Teng	15.89	–
Fu Ling	17.19	–
Tu Fu Ling	17.35	–
Gui Zhi	17.36	–
Lu Gen	17.77	–
Lian Qiao	17.93	–
Yan Hu Suo	18.15	–
Gou Teng	18.27	–
Ren Shen	18.40	–
Huo Ma Ren	18.53	–
(Huai) Niu Xi	19.03	–
Gu Sui Bu	19.47	–
Fu Xiao Mai	19.75	–
Suan Zao Ren	19.96	–
(Shi) Chang Pu	19.98	–
Gou Qi Zi	20.36	–
Bai Xian Pi	20.42	–
Shan Yu Rou	20.89	–
Xiang Fu	21.15	–
Gan Jiang	21.57	–
Ji Xue Teng	21.60	–
Yi Mu Cao	22.38	–
Fu Zi	22.38	–
Long Yan Rou	22.40	–
Fu Pen Zi	22.59	–
Jiao Gu Lan	22.65	–
Yi Yi Ren	23.11	–
Xie Bai	24.98	–
(Fen) Bi Xie	25.34	–
Chuan Lian Zi	25.66	–
Zhi Ke	25.87	–
Ma Huang	25.88	–
Ma Huang Gen	26.05	–
Dan Dou Chi	26.15	–
Dan Shen	26.16	–
Huang Qin	27.70	–
Sang Zhi	27.74	–
Chuan Niu Xi	27.77	–
Jiang Huang	27.94	–
Cang Er Zi	28.15	–
Wu Wei Zi	28.38	–
Ce Bai Ye	28.93	–
Hou Po	28.97	–
Yin Yang Huo	29.13	–
Huang Qi	29.18	–
Hong Jing Tian	29.25	–
Chuang Mu Xiang	29.43	–
Mao Dong Qing	29.63	–
Yu Jin	29.75	–
Xu Duan	30.15	–
Bai Zhi	30.61	–
Chai Hu	30.66	–
Mang Xiao	30.70	–
Gan Cao	30.74	–
She Gan	31.19	–
(Bai) Dou Kou	31.28	–
Sha Ren	31.82	–
Fu Shen	32.01	–
Wu Yao	32.07	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Che Qian Zi	32.61	–
Zhu Ling	32.82	–
Guang Huo Xiang	35.11	–
Zhi Gan Cao	35.47	–
Pu Gong Ying	35.98	–
Mi Huan Jun	36.19	–
Qiang Huo	36.37	–
E Zhu	36.97	–
Dong Gua Zi	37.21	–
Sha Shen (Bei)	37.32	–
Bai Zhu	37.63	–
Mai Men Dong	37.92	–
Bai Hua She She Cao	38.22	–
Chi Shao (Yao)	39.03	–
Fang Feng	40.45	–
Du Huo	41.00	–
Jing Jie	42.29	–
Sang Ye	42.38	–
Qin Jiao	42.40	–
Wang Bu Liu Xing	43.19	–
Zi Hua Di Ding	43.72	–
Huang Lian	44.07	–
Ze Lan	44.26	–
Ju Hua	44.89	–
Wu Zhu Yu	45.24	–
Bo He	45.41	–
Qing Pi	45.94	–
Mu Dan Pi	46.80	–
Chuan Mu Tong	46.92	–
Sang Ji Shend	46.93	–
Xin Yi	47.67	–
Ba Ji Tian	47.82	–
Wu Mei	48.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gua Lou* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62317	62317	0.00	14.11
62318	62318	0.00	14.20
62547	62547	0.00	5.28
62548	62548	0.00	4.37
62797	62797	0.00	4.29
62798	62798	0.00	6.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Guang Huo Xiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	61079-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Guang Huo Xiang; Pogostemonis herba

Special notes

When selecting the *Guang Huo Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Guang Huo Xiang	2	0	1

Second-stage model

For differentiation of the substance/substance group *Guang Huo Xiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Guang Huo Xiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Guang Huo Xiang	G007H2005822	62589	40	from supplier
PhytoComm	Guang Huo Xiang	G007H2005822	62590	40	from supplier
PhytoComm	Guang Huo Xiang	G007HS347TM1	63019	40	from supplier
PhytoComm	Guang Huo Xiang	G007HS347TM1	63020	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Guang Huo Xiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Guang Huo Xiang*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Guang Huo Xiang	G007H2005822	62589†	20
PhytoComm	Guang Huo Xiang	G007H2005822	62590†	20
PhytoComm	Guang Huo Xiang	G007HS347TM1	63019†	20
PhytoComm	Guang Huo Xiang	G007HS347TM1	63020†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Guang Huo Xiang*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Guang Huo Xiang	G007H2005521	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Guang Huo Xiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Guang Huo Xiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	159	1	24 439
Type B	0	76	4	12 397
Type C	0	0	2	855

The substance/substance group *Guang Huo Xiang* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9964 % (> 99.9814 %)	99.3750 % (> 97.5000 %)
Type B	100.0000 % (> 99.9403 %)	95.0000 % (> 91.2500 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†] These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Guang Huo Xiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ce Bai Ye	3.79	–
Dan Dou Chi	4.97	–
Yin Yang Huo	5.57	–
Bai Xian Pi	5.79	–
Tu Fu Ling	7.77	–
Hou Po	7.98	–
Ye Jiao Teng	8.14	–
Shen Qu	8.18	–
Ren Dong Teng	8.23	–
Ling Zhi	8.35	–
Jing Jie	8.59	–
Yu Jin	8.82	–
Chai Hu	9.18	–
Mao Dong Qing	9.31	–
Che Qian Zi	9.53	–
Zhu Ling	9.71	–
Shan Yao	9.94	–
Dan Shen	10.85	–
Fu Zi	10.87	–
Qiang Huo	10.99	–
Gu Sui Bu	11.23	–
Chuan Xiong	11.57	–
Fu Pen Zi	12.08	–
Lian Zi	12.10	–
(Fen) Bi Xie	12.15	–
She Gan	12.37	–
Pi Pa Ye	12.88	–
Yan Hu Suo	12.90	–
Ji Li	13.57	–
Zi Hua Di Ding	13.89	–
Jin Yin Hua	15.55	–
Ji Xue Teng	15.99	–
Sha Ren	16.11	–
Ma Huang	16.30	–
Gou Teng	16.48	–
Rou Gui	16.54	–
Gan Cao	16.60	–
Hong Jing Tian	16.61	–
Suan Zao Ren	16.63	–
He Huan Pi	16.98	–
Lian Qiao	17.03	–
Huo Ma Ren	17.06	–
Ma Huang Gen	17.46	–
Lu Gen	17.49	–
Mang Xiao	17.54	–
Tian Hua Fen	17.82	–
Bai Shao Yao	18.34	–
Bo He	18.40	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gua Lou	18.46	–
Fu Ling	19.32	–
Mu Zei	19.50	–
Ban Lan Gen	19.72	–
Gui Zhi	19.86	–
Jiao Gu Lan	20.29	–
Du Zhong	20.86	–
Sheng Jiang	21.17	–
Qing Pi	21.99	–
Huang Bai	22.63	–
Di Gu Pi	22.87	–
Yi Yi Ren	23.35	–
Zhe Bei Mu	23.38	–
Cang Zhu	23.53	–
Yuan Zhi	23.87	–
Zhi Gan Cao	23.97	–
Tao Ren	24.24	–
Fo Shou	24.42	–
Huang Lian	24.77	–
Zhi Ke	25.29	–
Wu Wei Zi	25.37	–
Ren Shen	26.70	–
Chen Pi	27.27	–
Ci Wu Jia	27.45	–
Tai Zi Shen	28.01	–
Gou Qi Zi	28.16	–
Bai Zi Ren	28.71	–
Ze Xie	28.98	–
Ban Xia (Jiang)	29.38	–
Fu Xiao Mai	29.99	–
Zhu Ru	30.31	–
Chuang Mu Xiang	30.55	–
Bai Jiang Cao	31.35	–
Yu Zhu	32.19	–
Sang Zhi	32.90	–
Qing Hao	32.90	–
Chi Shao (Yao)	33.89	–
Fu Shen	33.94	–
Dang Gui Wei	34.53	–
Jie Geng	34.93	–
Dang Gui	35.74	–
Huang Qin	36.14	–
Hong Hua	36.18	–
San Qi	36.51	–
Lai Fu Zi	36.96	–
Xie Bai	37.26	–
Jiu Da Huang	39.60	–
Long Yan Rou	39.66	–
Ban Zhi Lian	39.85	–
Pu Gong Ying	40.53	–
Shan Yu Rou	42.14	–
Cang Er Zi	43.00	–
Ku Shen	44.20	–
Hua Shi	45.03	–
Ze Lan	45.67	–
Bai Zhu	45.77	–
Sang Ye	47.52	–
Wu Mei	47.53	–
Zhi Mu	48.51	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	49.81	–
Shu Di (Huang)	51.77	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Guang Huo Xiang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62589	62589	0.00	17.68
62590	62590	0.00	17.54
63019	63019	0.00	4.97
63020	63020	0.00	3.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Gui Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60189-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Gui Zhi; Cinnamomi cassiae ramulus

Special notes

When selecting the *Gui Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Gui Zhi	3	0	4

Second-stage model

For differentiation of the substance/substance group *Gui Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Gui Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Gui Zhi	G070HS205SH1	62563	40	from supplier
PhytoComm	Gui Zhi	G070HS205SH1	62564	40	from supplier
PhytoComm	Gui Zhi	G070HS205TK1	62931	40	from supplier
PhytoComm	Gui Zhi	G070HS205TK1	62932	40	from supplier
PhytoComm	Gui Zhi	G070HS205TH1	62933	40	from supplier
PhytoComm	Gui Zhi	G070HS205TH1	62934	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Gui Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Gui Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Gui Zhi	G070HS205SH1	62563†	20
PhytoComm	Gui Zhi	G070HS205SH1	62564†	20
PhytoComm	Gui Zhi	G070HS205TK1	62931†	20
PhytoComm	Gui Zhi	G070HS205TK1	62932†	20
PhytoComm	Gui Zhi	G070HS205TH1	62933†	20
PhytoComm	Gui Zhi	G070HS205TH1	62934†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Gui Zhi*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Gui Zhi	G070H1001221	2
Phytocomm	Gui Zhi	g070h1001422	1
Phytocomm	Gui Zhi	G070H1001523	2
Phytocomm	Gui Zhi	G070H100422	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Gui Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Gui Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	239	1	24 359
Type B	1	112	8	12 356
Type C	1	0	6	850

The substance/substance group *Gui Zhi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9964 % (> 99.9814 %)	99.5833 % (> 98.3333 %)
Type B	99.9952 % (> 99.9654 %)	93.3333 % (> 90.8333 %)
Type C	99.7674 % (> 99.1804 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Gui Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Rou Gui	4.94	–
Sheng Jiang	4.99	–
He Huan Pi	6.20	–
Ban Xia (Jiang)	7.02	–
Mu Zei	7.22	–
Gou Teng	7.63	–
Yi Yi Ren	7.65	–
Ma Huang Gen	7.72	–
Fu Ling	8.08	–
Tao Ren	8.17	–
Ci Wu Jia	8.80	–
Ji Xue Teng	8.80	–
Di Gu Pi	8.88	–
Zhu Ru	9.34	–
Lu Gen	9.53	–
Tai Zi Shen	10.04	–
Lai Fu Zi	10.15	–
Fu Xiao Mai	10.24	–
Fu Shen	10.81	–
Gu Sui Bu	12.28	–
Bai Zi Ren	13.01	–
Ji Li	13.71	–
Fo Shou	13.72	–
Lian Zi	13.85	–
Huo Ma Ren	13.89	–
Bai Xian Pi	14.17	–
Bai Shao Yao	14.20	–
Ye Jiao Teng	14.39	–
Ling Zhi	14.44	–
Zhe Bei Mu	15.67	–
Shen Qu	15.93	–
Ren Dong Teng	15.94	–
(Fen) Bi Xie	16.17	–
Gua Lou	18.25	–
Suan Zao Ren	20.03	–
Chen Pi	20.14	–
Tu Fu Ling	20.99	–
Ze Xie	21.10	–
Tian Hua Fen	21.21	–
She Gan	21.57	–
Zhu Ling	21.90	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	22.36	–
Chuan Xiong	23.34	–
Ren Shen	23.73	–
Dan Shen	24.03	–
Yuan Zhi	24.38	–
Dan Dou Chi	24.75	–
Yu Jin	24.80	–
Yu Zhu	25.16	–
Lian Qiao	26.41	–
Pi Pa Ye	26.65	–
Dang Gui	27.53	–
Yan Hu Suo	27.71	–
Ban Lan Gen	27.74	–
Shan Yao	27.88	–
Ce Bai Ye	27.98	–
Ma Huang	28.17	–
Yin Yang Huo	28.44	–
Hong Jing Tian	28.59	–
Cang Zhu	28.60	–
Jie Geng	29.69	–
Zhi Mu	29.86	–
Dang Gui Wei	30.75	–
Jiao Gu Lan	30.87	–
Zhi Ke	31.28	–
Mang Xiao	31.31	–
Long Yan Rou	31.77	–
Mao Dong Qing	31.84	–
Chai Hu	32.04	–
Fu Pen Zi	32.18	–
Guang Huo Xiang	32.46	–
Huang Qin	33.46	–
Fu Zi	34.87	–
San Qi	36.01	–
Mu Gua	36.31	–
Gan Cao	36.79	–
Shan Yu Rou	36.83	–
Qiang Huo	38.41	–
Hou Po	38.63	–
Sang Zhi	38.87	–
Jing Jie	39.04	–
Wu Wei Zi	41.42	–
Zi Hua Di Ding	43.14	–
Gou Qi Zi	43.23	–
Huang Lian	43.69	–
Sha Ren	44.15	–
Chuang Mu Xiang	44.85	–
Che Qian Zi	45.71	–
Chi Shao (Yao)	46.90	–
Ku Shen	47.53	–
Bai Zhu	50.06	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Gui Zhi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62563	62563	0.00	4.99
62564	62564	0.00	4.99
62931	62931	0.00	5.42
62932	62932	0.00	4.94
62933	62933	0.00	6.25
62934	62934	0.00	5.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Han Lian Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60540-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Han Lian Cao; Ecliptae prostratae herba

Special notes

When selecting the *Han Lian Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Han Lian Cao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Han Lian Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Han Lian Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Han Lian Cao	G095HS141RW1	62679	40	from supplier
PhytoComm	Han Lian Cao	G095HS141RW1	62680	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Han Lian Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Han Lian Cao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Han Lian Cao	G095HS141RW1	62679†	20
PhytoComm	Han Lian Cao	G095HS141RW1	62680†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Han Lian Cao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Han Lian Cao	g095h0770321	1
Phytocomm	Han Lian Cao	G095H0770321	1
PhytoComm	Han Lian Cao	G095H0770321	2
Phytocomm	Han Lian Cao	G095H0770521	4

- 849 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Han Lian Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Han Lian Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	8	849

The substance/substance group *Han Lian Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8253 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Han Lian Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Hua She She Cao	6.53	–
Yu Xing Cao	10.43	–
Xi Xian Cao	13.34	–
Xin Yi	13.96	–
Pu Gong Ying	15.01	–
Jiang Huang	15.91	–
Xia Ku Cao	16.02	–
Zhi Shi	16.07	–
Tu Fu Ling	16.98	–
Sang Ye	18.59	–
Huang Lian	18.81	–
Hu Zhang	19.53	–
Shan Yao	21.43	–
Ze Lan	21.60	–
Nü Zhen Zi	23.54	–
Sha Ren	25.65	–
He Huan Pi	27.38	–
Xiang Fu	27.81	–
Mang Xiao	28.59	–
Yi Mu Cao	29.79	–
Rou Cong Rong	32.82	–
Yin Chen Hao	34.02	–
Du Zhong	34.36	–
Gu Sui Bu	35.81	–
Qiang Huo	37.48	–
Sang Ji Shend	39.44	–
Yu Jin	41.18	–
Fu Zi	41.32	–
Yan Hu Suo	41.96	–
Gou Teng	42.76	–
Gua Lou	43.58	–
Ji Li	44.02	–
Cang Er Zi	44.49	–
Niu Bang Zi	44.91	–
E Zhu	45.44	–
Hou Po	45.94	–
Dan Shen	46.34	–
Chen Pi	46.70	–
Bo He	46.94	–
Wu Mei	47.02	–
(Bai) Dou Kou	47.58	–
Che Qian Zi	48.64	–
Bai Jiang Cao	49.06	–
Ju Hua	49.44	–
Wu Yao	49.79	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Han Lian Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62679	62679	0.00	6.53
62680	62680	0.00	9.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**He Huan Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60175-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

He Huan Pi; Albiziae cortex

Special notes

When selecting the *He Huan Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
He Huan Pi	4	0	2

Second-stage model

For differentiation of the substance/substance group *He Huan Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *He Huan Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	He Huan Pi	G010H0635822	62402	40	from supplier
PhytoComm	He Huan Pi	G010H0635822	62403	40	from supplier
PhytoComm	He Huan Pi	G010HS114SK1	62645	40	from supplier
PhytoComm	He Huan Pi	G010HS114SK1	62646	40	from supplier
PhytoComm	He Huan Pi	G010HS114SW1	62949	40	from supplier
PhytoComm	He Huan Pi	G010HS114SW1	62950	40	from supplier
PhytoComm	He Huan Pi	G010HS114TK1	62951	40	from supplier
PhytoComm	He Huan Pi	G010HS114TK1	62952	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 320 spectra of 8 reference samples from the substance/substance group *He Huan Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 4 different batches.
- 24 280 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 160 spectra of 8 reference samples from the substance/substance group *He Huan Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	He Huan Pi	G010H0635822	62402†	20
PhytoComm	He Huan Pi	G010H0635822	62403†	20
PhytoComm	He Huan Pi	G010HS114SK1	62645†	20
PhytoComm	He Huan Pi	G010HS114SK1	62646†	20

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	He Huan Pi	G010HS114SW1	62949†	20
PhytoComm	He Huan Pi	G010HS114SW1	62950†	20
PhytoComm	He Huan Pi	G010HS114TK1	62951†	20
PhytoComm	He Huan Pi	G010HS114TK1	62952†	20

- 12 317 spectra from a total of 305 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *He Huan Pi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	He Huan Pi	g010h0635221	1
Phytocomm	He Huan Pi	G010H0635423	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *He Huan Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *He Huan Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	17	305	15	24 263
Type B	27	140	20	12 290
Type C	0	1	1	855

The substance/substance group *He Huan Pi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

†These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

	Specificity	Recognition rate
Type A	99.8917 % (> 99.8766 %)	95.3125 % (> 94.3750 %)
Type B	99.6667 % (> 99.6368 %)	87.5000 % (> 85.6250 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *He Huan Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Rou Gui	2.65	–
Tao Ren	3.64	–
Huo Ma Ren	4.42	–
Sheng Jiang	4.70	–
Ci Wu Jia	4.72	–
Mu Zei	5.02	–
Ma Huang Gen	5.06	–
Gui Zhi	5.20	–
Gou Teng	5.64	–
Lai Fu Zi	6.15	–
Di Gu Pi	7.23	–
Bai Zi Ren	7.30	–
Ji Xue Teng	7.72	–
Lian Zi	8.09	–
Bai Shao Yao	8.15	–
Tai Zi Shen	8.17	–
Fu Xiao Mai	8.42	–
Gu Sui Bu	8.79	–
Fu Ling	8.87	–
Lu Gen	9.72	–
Yi Yi Ren	9.90	–
Ban Xia (Jiang)	9.92	–
Ling Zhi	10.14	–
Ji Li	10.48	–
Shen Qu	11.17	–
Zhu Ru	11.27	–
Zhe Bei Mu	11.48	–
Bai Xian Pi	11.56	–
Fu Shen	11.77	–
Ye Jiao Teng	12.53	–
Tian Hua Fen	13.32	–
Fo Shou	13.42	–
Suan Zao Ren	14.00	–
(Fen) Bi Xie	14.34	–
Ze Xie	14.37	–
Ren Dong Teng	15.25	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiang Huang	15.46	–
Zhu Ling	15.86	–
Gua Lou	16.25	–
She Gan	16.79	–
Tu Fu Ling	17.28	–
Chuan Xiong	17.71	–
Yu Jin	19.02	–
Ce Bai Ye	19.41	–
Lian Qiao	19.92	–
Yuan Zhi	20.00	–
Pi Pa Ye	20.20	–
Shan Yao	20.48	–
Chen Pi	20.59	–
Wu Yao	20.63	–
(Bai) Dou Kou	20.71	–
Dan Dou Chi	21.50	–
Sha Ren	21.61	–
Yan Hu Suo	22.73	–
Dan Shen	22.73	–
Jin Yin Hua	23.07	–
Mao Dong Qing	25.31	–
Fu Pen Zi	25.35	–
Chai Hu	25.51	–
Fu Zi	26.15	–
Cang Zhu	26.16	–
Yin Yang Huo	26.56	–
Jiao Gu Lan	26.73	–
Sang Bai Pi	27.06	–
Yi Mu Cao	27.12	–
Yu Zhu	27.31	–
Ban Lan Gen	27.55	–
Yin Chen Hao	27.84	–
Hong Jing Tian	28.12	–
Mang Xiao	28.38	–
Guang Huo Xiang	28.53	–
Dang Gui Wei	28.63	–
Jie Geng	29.36	–
Zhi Ke	30.05	–
Ma Huang	30.16	–
San Qi	30.21	–
Hou Po	30.26	–
Dang Gui	30.43	–
Zhi Mu	31.05	–
Ren Shen	31.57	–
E Zhu	31.82	–
Sang Ji Shend	32.35	–
Gan Cao	32.35	–
Che Qian Zi	32.80	–
He Shou Wu	33.06	–
Shan Yu Rou	33.78	–
Qiang Huo	33.91	–
(Shi) Chang Pu	33.94	–
Bu Gu Zhi	34.57	–
Huang Qin	36.41	–
Dan Zhu Ye	36.82	–
Tu Si Zi	37.82	–
Hu Zhang	37.94	–
Jing Jie	37.95	–
Wu Wei Zi	38.88	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Zhu Yu	39.04	–
Mu Gua	39.50	–
Zi Hua Di Ding	39.90	–
Bai Hua She She Cao	39.96	–
Niu Bang Zi	40.14	–
Long Yan Rou	40.77	–
Chuang Mu Xiang	41.12	–
Gou Qi Zi	41.32	–
Huang Lian	41.45	–
Xia Ku Cao	41.45	–
Xiao Hui Xiang	41.87	–
Sang Zhi	42.31	–
Xin Yi	42.64	–
Chuan Niu Xi	42.97	–
Ze Lan	42.97	–
Ju Hua	43.33	–
Mu Dan Pi	43.49	–
Xiang Fu	43.96	–
Du Zhong	45.08	–
Ku Shen	45.11	–
Xi Xian Cao	45.67	–
Ge Gen	45.70	–
Nü Zhen Zi	45.88	–
Qing Pi	46.15	–
Fang Feng	46.58	–
Zhi Gan Cao	47.23	–
Bo He	47.67	–
Xie Bai	47.91	–
Xuan Fu Hua	49.86	–
Dong Gua Zi	50.77	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *He Huan Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62402	62402	0.00	15.49
62403	62403	0.00	15.46
62645	62645	0.00	3.64
62646	62646	0.00	4.70
62949	62949	0.00	4.32
62950	62950	0.00	4.42
62951	62951	0.00	2.77
62952	62952	0.00	2.65

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**He Shou Wu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60101-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

He Shou Wu; Polygoni multiflori radix

Special notes

When selecting the *He Shou Wu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
He Shou Wu	2	0	2

Second-stage model

For differentiation of the substance/substance group *He Shou Wu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *He Shou Wu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	He Shou Wu	G198H0718821	62301	40	from supplier
PhytoComm	He Shou Wu	G198H0718821	62302	40	from supplier
PhytoComm	He Shou Wu	G198H0718921	62745	40	from supplier
PhytoComm	He Shou Wu	G198H0718921	62746	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *He Shou Wu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *He Shou Wu*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	He Shou Wu	G198H0718821	62301†	20
PhytoComm	He Shou Wu	G198H0718821	62302†	20
PhytoComm	He Shou Wu	G198H0718921	62745†	20
PhytoComm	He Shou Wu	G198H0718921	62746†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 7 spectra from 6 *Apo-Ident* customers from 3 batches from the substance/substance group *He Shou Wu*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	He Shou Wu	g198h0718222	1
Phytocomm	He Shou Wu	G198H0718222	1
PhytoComm	He Shou Wu	G198H0718222	1
Phytocomm	He Shou Wu	G198H0718421	3
PhytoComm	He Shou Wu	G198H0718421	1

- 850 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *He Shou Wu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *He Shou Wu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	2	0	7	848

The substance/substance group *He Shou Wu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.7674 % (> 99.1802 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *He Shou Wu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Xuan Fu Hua	6.23	–
Sang Ji Shend	10.25	–
Yin Chen Hao	12.71	–
Mu Dan Pi	14.41	–
Gu Sui Bu	14.57	–
Jing Jie	15.29	–
He Huan Pi	16.00	–
Dan Zhu Ye	18.07	–
Sang Bai Pi	20.64	–
Chen Pi	23.02	–
Sha Ren	23.30	–
Mang Xiao	23.98	–
Jin Qian Cao	25.94	–
Ge Gen	26.30	–
Shan Yu Rou	26.92	–
Ji Li	28.08	–
Du Zhong	28.79	–
Wu Yao	29.02	–
Jiang Huang	29.44	–
Hong Jing Tian	29.66	–
Nü Zhen Zi	29.81	–
Bu Gu Zhi	29.89	–
Wu Zhu Yu	30.62	–
Gou Teng	31.70	–
Chuan Niu Xi	34.51	–
Dang Gui	34.89	–
Tu Si Zi	35.29	–
(Shi) Chang Pu	36.09	–
(Bai) Dou Kou	36.22	–
Shan Yao	36.86	–
Bai Hua She She Cao	37.14	–
E Zhu	38.14	–
Fang Feng	39.07	–
Wu Jia Pi	39.32	–
Yi Mu Cao	39.76	–
Sang Ye	40.37	–
Ding Xiang	40.52	–
Zhi Shi	40.77	–
Ju Hua	41.40	–
Xiang Fu	42.08	–
Bing Lang	42.15	–
Hu Zhang	42.39	–
Huang Qi	42.52	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Di Gu Pi	42.63	–
Xiao Hui Xiang	44.90	–
Ba Ji Tian	47.44	–
Chuan Lian Zi	47.49	–
Zhi Ke	48.19	–
Xia Ku Cao	48.27	–
Du Huo	48.71	–
Xin Yi	48.93	–
Gua Lou	49.64	–
Ku Shen	49.98	–
Hua Shi	50.23	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *He Shou Wu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62301	62301	0.00	6.63
62302	62302	0.00	6.23
62745	62745	0.00	14.63
62746	62746	0.00	14.41

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Hong Hua**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60206-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Hong Hua; Carthami flos

Special notes

When selecting the *Hong Hua* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Hong Hua	2	0	2

Second-stage model

For differentiation of the substance/substance group *Hong Hua* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Hong Hua*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Hong Hua	G057HS173SG1	62663	40	from supplier
PhytoComm	Hong Hua	G057HS173SG1	62664	40	from supplier
PhytoComm	Hong Hua	G057HS173TG1	62915	40	from supplier
PhytoComm	Hong Hua	G057HS173TG1	62916	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Hong Hua*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Hong Hua*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Hong Hua	G057HS173SG1	62663†	20
PhytoComm	Hong Hua	G057HS173SG1	62664†	20
PhytoComm	Hong Hua	G057HS173TG1	62915†	20
PhytoComm	Hong Hua	G057HS173TG1	62916†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Hong Hua*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Hong Hua	g057h0901121	1
Phytocomm	Hong Hua	G057H0901221	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Hong Hua* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hong Hua* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	2	855

The substance/substance group *Hong Hua* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hong Hua* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shan Yao	10.92	–
Zhi Gan Cao	11.05	–
Jiao Gu Lan	11.80	–
Zhi Ke	12.48	–
Wu Wei Zi	14.91	–
Huang Bai	15.14	–
Gan Cao	15.38	–
Chuang Mu Xiang	15.61	–
Jin Yin Hua	15.68	–
Qing Hao	15.84	–
Zi Hua Di Ding	17.21	–
Qing Pi	17.31	–
Ma Huang	17.46	–
Shan Yu Rou	18.33	–
Hong Jing Tian	18.71	–
Sang Ye	18.83	–
Ren Shen	19.96	–
Dan Dou Chi	21.50	–
Suan Zao Ren	21.79	–
Dan Shen	22.01	–
Huang Lian	22.06	–
Bo He	22.63	–
Sha Ren	22.81	–
Hou Po	23.01	–
Gou Qi Zi	23.03	–
Mao Dong Qing	23.36	–
Xie Bai	23.43	–
Wu Mei	23.44	–
Cang Zhu	23.71	–
Ban Lan Gen	23.75	–
Chai Hu	24.07	–
Zhe Bei Mu	24.40	–
Fu Zi	25.13	–
Chuan Xiong	25.43	–
Qiang Huo	25.52	–
Fu Pen Zi	25.74	–
Jie Geng	26.40	–
Lian Qiao	26.63	–
Guang Huo Xiang	26.90	–
Ku Shen	27.11	–
Chen Pi	27.11	–
Dang Gui	27.13	–
Gua Lou	27.35	–
Ye Jiao Teng	27.42	–
Pu Gong Ying	27.70	–
Tian Hua Fen	27.73	–
Ji Li	27.86	–
Du Zhong	27.91	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yin Yang Huo	28.27	–
Cang Er Zi	28.50	–
Huang Qin	28.58	–
Mang Xiao	29.40	–
She Gan	29.67	–
Ling Zhi	30.09	–
Ren Dong Teng	30.33	–
Chi Shao (Yao)	30.74	–
Jing Jie	30.74	–
Lu Gen	32.20	–
Mu Zei	32.22	–
Gu Sui Bu	32.52	–
Pi Pa Ye	32.57	–
Bai Jiang Cao	32.95	–
Yan Hu Suo	33.29	–
He Huan Pi	34.32	–
Bai Shao Yao	34.34	–
Ze Xie	35.10	–
Tu Fu Ling	35.11	–
Yu Zhu	35.51	–
San Qi	36.11	–
Lian Zi	36.92	–
Zhi Mu	37.15	–
Yuan Zhi	37.46	–
Niu Bang Zi	37.56	–
Ze Lan	37.93	–
Di Gu Pi	38.00	–
Zi Su Zi	38.73	–
Ce Bai Ye	38.85	–
Long Yan Rou	38.85	–
Bai Zi Ren	38.94	–
Mu Gua	39.04	–
Tao Ren	39.07	–
Tai Zi Shen	40.20	–
Xi Xian Cao	40.93	–
Dang Gui Wei	41.51	–
Bai Xian Pi	41.68	–
(Fen) Bi Xie	41.86	–
Che Qian Zi	41.88	–
Sheng Jiang	42.25	–
Huo Ma Ren	42.98	–
Yu Jin	43.23	–
Sang Zhi	44.71	–
Ji Xue Teng	44.77	–
Gou Teng	45.28	–
Zhu Ling	45.35	–
Xiang Fu	45.54	–
Shen Qu	45.84	–
Rou Gui	46.47	–
Nü Zhen Zi	46.55	–
Zhu Ru	47.21	–
Lai Fu Zi	47.47	–
Yi Mu Cao	47.94	–
Gui Zhi	47.95	–
Ban Zhi Lian	48.77	–
Fu Ling	49.52	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/substance group *Hong Hua* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62663	62663	0.00	11.13
62664	62664	0.00	10.92
62915	62915	0.00	11.05
62916	62916	0.00	12.71

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Hong Jing Tian**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60031-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Hong Jing Tian; Rhodiolae crenulatae radix

Special notes

When selecting the *Hong Jing Tian* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Hong Jing Tian	2	0	2

Second-stage model

For differentiation of the substance/substance group *Hong Jing Tian* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Hong Jing Tian*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Hong Jing Tian	G237H0906821	62305	40	from supplier
PhytoComm	Hong Jing Tian	G237H0906821	62306	40	from supplier
PhytoComm	Hong Jing Tian	G237HS429SK1	62639	40	from supplier
PhytoComm	Hong Jing Tian	G237HS429SK1	62640	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Hong Jing Tian*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Hong Jing Tian*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Hong Jing Tian	G237H0906821	62305†	20
PhytoComm	Hong Jing Tian	G237H0906821	62306†	20
PhytoComm	Hong Jing Tian	G237HS429SK1	62639†	20
PhytoComm	Hong Jing Tian	G237HS429SK1	62640†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 1 *Apo-Ident* customers from 2 batches from the substance/substance group *Hong Jing Tian*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Hong Jing Tian	G237H0906221	1
Phytocomm	Hong Jing Tian	G237H0906321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Hong Jing Tian* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hong Jing Tian* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	78	2	12 397
Type C	0	0	2	855

The substance/substance group *Hong Jing Tian* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	97.5000 % (> 93.7500 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hong Jing Tian* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	4.73	–
Zhi Gan Cao	5.67	–
Dan Shen	8.65	–
Chuan Xiong	9.23	–
Dan Dou Chi	10.14	–
Zi Hua Di Ding	10.23	–
Gan Cao	10.92	–
Hou Po	11.58	–
Guang Huo Xiang	11.94	–
Gua Lou	12.27	–
Fu Pen Zi	12.60	–
Gou Qi Zi	12.81	–
Ban Lan Gen	13.32	–
Pi Pa Ye	13.33	–
Fu Zi	14.13	–
Gu Sui Bu	14.19	–
Ren Dong Teng	14.65	–
Tian Hua Fen	14.92	–
Jiao Gu Lan	15.19	–
Zhi Ke	15.27	–
Ma Huang	15.33	–
Shan Yao	15.46	–
Suan Zao Ren	15.46	–
Chuang Mu Xiang	15.72	–
She Gan	15.84	–
Cang Zhu	15.99	–
Yin Yang Huo	16.52	–
Chai Hu	16.60	–
Bai Shao Yao	16.99	–
Yuan Zhi	17.44	–
Ye Jiao Teng	17.49	–
Qiang Huo	17.61	–
Lian Qiao	17.94	–
Mu Zei	18.11	–
Chen Pi	18.25	–
Yan Hu Suo	18.56	–
Lian Zi	18.57	–
Wu Wei Zi	18.69	–
Hong Hua	18.78	–
Shen Qu	18.81	–
Mao Dong Qing	19.49	–
Bo He	19.50	–
Ji Li	19.68	–
Lu Gen	19.70	–
Ze Xie	20.43	–
Qing Pi	20.53	–
Sang Ji Shend	20.70	–
Ce Bai Ye	21.23	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Shen	21.53	–
Bai Xian Pi	21.64	–
Huang Bai	22.08	–
Zhe Bei Mu	22.60	–
Tu Fu Ling	22.71	–
Jing Jie	22.77	–
He Huan Pi	22.80	–
Sha Ren	22.82	–
Di Gu Pi	23.03	–
Che Qian Zi	23.08	–
Mang Xiao	23.31	–
Ji Xue Teng	23.36	–
Yu Jin	23.37	–
Dang Gui	23.52	–
Ku Shen	23.64	–
Jie Geng	24.19	–
Ling Zhi	25.02	–
Long Yan Rou	25.07	–
Xie Bai	25.44	–
Gou Teng	25.45	–
Nü Zhen Zi	25.54	–
Yu Zhu	25.78	–
(Fen) Bi Xie	26.03	–
Tao Ren	26.15	–
Huang Lian	26.94	–
Ge Gen	26.97	–
Dang Gui Wei	27.01	–
Tai Zi Shen	27.21	–
Shan Yu Rou	27.23	–
Fang Feng	27.32	–
Zhu Ling	27.37	–
Huang Qin	27.41	–
Du Zhong	27.44	–
Yi Mu Cao	27.54	–
Xuan Fu Hua	27.71	–
Sheng Jiang	27.72	–
Gui Zhi	27.83	–
Chi Shao (Yao)	28.15	–
Huo Ma Ren	28.75	–
San Qi	28.92	–
Fu Ling	29.52	–
He Shou Wu	29.75	–
Bai Zi Ren	29.79	–
Rou Gui	30.82	–
Ci Wu Jia	31.41	–
Cang Er Zi	31.62	–
Sang Bai Pi	32.44	–
Bai Jiang Cao	33.31	–
Sang Zhi	33.58	–
Ma Huang Gen	34.51	–
Lai Fu Zi	34.51	–
Qing Hao	35.16	–
Zhi Mu	35.76	–
Wu Zhu Yu	36.59	–
Sang Ye	36.85	–
Mu Gua	37.16	–
Ban Xia (Jiang)	38.31	–
Dan Zhu Ye	38.43	–
Yin Chen Hao	38.90	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fo Shou	39.70	–
Zhu Ru	39.84	–
Hu Zhang	40.27	–
Pu Gong Ying	42.77	–
Yi Yi Ren	43.00	–
Bu Gu Zhi	43.21	–
Fu Xiao Mai	43.95	–
Ze Lan	44.50	–
Mu Dan Pi	45.68	–
(Shi) Chang Pu	45.88	–
Ban Zhi Lian	46.17	–
Xia Ku Cao	47.71	–
Ju Hua	47.80	–
Niu Bang Zi	49.43	–
Fu Shen	49.45	–
Huang Qi	49.86	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hong Jing Tian* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62305	62305	0.00	20.72
62306	62306	0.00	20.70
62639	62639	0.00	4.95
62640	62640	0.00	4.73

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Hou Po**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50289-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Hou Po; Magnoliae officinalis cortex

Special notes

When selecting the *Hou Po* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Hou Po	1	0	1

Second-stage model

For differentiation of the substance/substance group *Hou Po* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Hou Po*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Hou Po	G154HS184SH1	62613	40	from supplier
PhytoComm	Hou Po	G154HS184SH1	62614	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Hou Po*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Hou Po*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Hou Po	G154HS184SH1	62613†	20
PhytoComm	Hou Po	G154HS184SH1	62614†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Hou Po*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Hou Po	G154H0917222	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Hou Po* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hou Po* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	2	40	0	12 435
Type C	0	0	1	856

The substance/substance group *Hou Po* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9810 % (> 99.9512 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hou Po* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chai Hu	7.08	–
Dan Dou Chi	7.17	–
Yu Jin	7.31	–
Yan Hu Suo	7.46	–
Mao Dong Qing	7.67	–
Ren Dong Teng	7.86	–
Ce Bai Ye	7.87	–
Shen Qu	7.87	–
Ling Zhi	8.12	–
Che Qian Zi	8.20	–
Shan Yao	8.50	–
Fu Pen Zi	8.50	–
Yin Yang Huo	8.65	–
Fu Zi	8.69	–
Bai Xian Pi	9.25	–
Guang Huo Xiang	9.38	–
Qiang Huo	9.58	–
Gu Sui Bu	9.75	–
Pi Pa Ye	9.88	–
Zhu Ling	10.28	–
Jing Jie	10.34	–
Chuan Xiong	10.59	–
Dan Shen	10.89	–
Ji Li	11.97	–
She Gan	12.45	–
Tu Fu Ling	12.72	–
Zi Hua Di Ding	13.00	–
Tian Hua Fen	13.33	–
Lian Zi	14.26	–
Ye Jiao Teng	14.40	–
Hong Jing Tian	14.87	–
Lian Qiao	15.22	–
Gan Cao	15.22	–
Jin Yin Hua	15.85	–
Qing Pi	15.85	–
Du Zhong	16.69	–
Ma Huang	17.05	–
Fu Ling	17.26	–
Jiao Gu Lan	17.99	–
Bo He	18.55	–
Ban Lan Gen	18.68	–
Gua Lou	18.88	–
Bai Shao Yao	18.96	–
Mu Zei	18.97	–
Ji Xue Teng	19.01	–
Suan Zao Ren	19.14	–
Sha Ren	19.23	–
(Fen) Bi Xie	19.59	–
He Huan Pi	19.61	–
Lu Gen	20.93	–
Qing Hao	20.95	–
Gou Teng	21.04	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang Gen	21.05	–
Zhe Bei Mu	21.13	–
Sheng Jiang	21.17	–
Huang Bai	21.25	–
Zhi Gan Cao	21.30	–
Huo Ma Ren	22.20	–
Gui Zhi	23.02	–
Wu Wei Zi	23.32	–
Rou Gui	23.48	–
Huang Lian	23.95	–
Cang Zhu	24.01	–
Ze Xie	24.59	–
Yuan Zhi	24.67	–
Ren Shen	24.82	–
Di Gu Pi	24.89	–
Zhi Ke	25.18	–
Bai Zi Ren	25.29	–
Yi Yi Ren	25.57	–
Tao Ren	27.32	–
Gou Qi Zi	27.38	–
Chen Pi	27.53	–
Tai Zi Shen	27.60	–
Bai Jiang Cao	29.14	–
Chuang Mu Xiang	31.26	–
Yu Zhu	31.87	–
Mang Xiao	32.54	–
Ban Xia (Jiang)	33.00	–
Zhu Ru	33.28	–
Fu Xiao Mai	34.72	–
Dang Gui	34.76	–
Ci Wu Jia	34.81	–
Jie Geng	35.11	–
Xie Bai	35.43	–
Huang Qin	35.66	–
Hong Hua	35.94	–
San Qi	36.57	–
Dang Gui Wei	37.02	–
Chi Shao (Yao)	37.39	–
Sang Zhi	37.49	–
Ku Shen	37.95	–
Fo Shou	38.13	–
Lai Fu Zi	38.13	–
Pu Gong Ying	39.04	–
Shan Yu Rou	39.97	–
Cang Er Zi	40.09	–
Fu Shen	41.21	–
Long Yan Rou	41.54	–
Sang Ye	42.51	–
Ze Lan	43.86	–
Ban Zhi Lian	46.50	–
Xi Xian Cao	46.69	–
Zhi Mu	48.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hou Po* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62613	62613	0.00	7.17
62614	62614	0.00	7.08

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Hu Zhang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60130-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Hu Zhang; Polygoni cuspidati rhizoma et radix

Special notes

When selecting the *Hu Zhang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Hu Zhang	1	0	1

Second-stage model

For differentiation of the substance/substance group *Hu Zhang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Hu Zhang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Hu Zhang	G197HS373SK1	62565	40	from supplier
PhytoComm	Hu Zhang	G197HS373SK1	62566	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Hu Zhang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Hu Zhang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Hu Zhang	G197HS373SK1	62565†	20
PhytoComm	Hu Zhang	G197HS373SK1	62566†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Hu Zhang*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Hu Zhang	g197h0851321	1
Phytocomm	Hu Zhang	G197H0851321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Hu Zhang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hu Zhang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	2	855

The substance/substance group *Hu Zhang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hu Zhang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zhi Shi	10.84	–
Han Lian Cao	11.47	–
Yu Xing Cao	13.22	–
Bai Hua She She Cao	16.07	–
Sang Ye	16.20	–
Xi Xian Cao	18.29	–
Xin Yi	18.37	–
Huang Lian	21.95	–
Xia Ku Cao	24.37	–
Ze Lan	25.05	–
Pu Gong Ying	25.25	–
Sha Ren	26.20	–
Mang Xiao	26.52	–
Jiang Huang	28.48	–
Yi Mu Cao	29.25	–
Shan Yao	31.12	–
Nü Zhen Zi	32.28	–
Xiang Fu	32.40	–
Tu Fu Ling	32.93	–
Sang Ji Shend	34.48	–
Yin Chen Hao	36.27	–
He Huan Pi	36.97	–
Rou Cong Rong	37.55	–
Gu Sui Bu	38.54	–
Gou Teng	41.20	–
Bo He	47.13	–
Qiang Huo	48.39	–
Ju Hua	49.00	–
Du Zhong	49.76	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hu Zhang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62565	62565	0.00	10.84
62566	62566	0.00	10.96

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Hua Shi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50308-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Hua Shi; Talcum

Special notes

When selecting the *Hua Shi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Hua Shi	2	0	1

Second-stage model

For differentiation of the substance/substance group *Hua Shi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Hua Shi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Hua Shi	G238HS289RM1	62425	40	from supplier
PhytoComm	Hua Shi	G238HS289RM1	62426	40	from supplier
PhytoComm	Hua Shi	G238HS289SK1	62637	40	from supplier
PhytoComm	Hua Shi	G238HS289SK1	62638	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Hua Shi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Hua Shi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Hua Shi	G238HS289RM1	62425†	20
PhytoComm	Hua Shi	G238HS289RM1	62426†	20
PhytoComm	Hua Shi	G238HS289SK1	62637†	20
PhytoComm	Hua Shi	G238HS289SK1	62638†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Hua Shi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Hua Shi	g238h1332121	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Hua Shi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Hua Shi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	64	96	24 440
Type B	0	29	51	12 397
Type C	0	0	1	856

The substance/substance group *Hua Shi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	40.0000 % (> 38.1250 %)
Type B	100.0000 % (> 99.9403 %)	36.2500 % (> 32.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Hua Shi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	113.53	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Hua Shi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62425	62425	0.00	113.53
62426	62426	0.00	115.53
62637	62637	0.00	147.47
62638	62638	0.00	145.42

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Huang Bai**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60949-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Huang Bai; Phellodendri chinensis cortex

Special notes

When selecting the *Huang Bai* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Huang Bai	2	0	0

Second-stage model

For differentiation of the substance/substance group *Huang Bai* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Huang Bai*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Bai	G188HS262SH1	62697	40	from supplier
PhytoComm	Huang Bai	G188HS262SH1	62698	40	from supplier
PhytoComm	Huang Bai	G188HS262SV1	62837	40	from supplier
PhytoComm	Huang Bai	G188HS262SV1	62838	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Huang Bai*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Huang Bai*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Huang Bai	G188HS262SH1	62697†	20
PhytoComm	Huang Bai	G188HS262SH1	62698†	20
PhytoComm	Huang Bai	G188HS262SV1	62837†	20
PhytoComm	Huang Bai	G188HS262SV1	62838†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Huang Bai*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Huang Bai* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Huang Bai* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	0	857

The substance/substance group *Huang Bai* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Huang Bai* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zhi Ke	7.03	–
Bo He	7.35	–
Qing Hao	8.01	–
Qing Pi	8.59	–
Jiao Gu Lan	8.61	–
Fu Zi	10.11	–
Huang Lian	10.12	–
Zi Hua Di Ding	10.14	–
Dan Shen	10.38	–
Chai Hu	11.49	–
Qiang Huo	11.60	–
Tu Fu Ling	12.67	–
Hong Jing Tian	13.51	–
Sha Ren	13.68	–
Gan Cao	14.40	–
Du Zhong	14.52	–
Jing Jie	15.57	–
Bai Jiang Cao	15.61	–
Fu Pen Zi	15.88	–
Yin Yang Huo	16.31	–
Ye Jiao Teng	16.39	–
Yu Jin	16.51	–
Jin Yin Hua	16.93	–
Shan Yao	17.06	–
Hou Po	17.09	–
Sang Ye	17.72	–
Mao Dong Qing	17.75	–
Ma Huang	18.80	–
Yan Hu Suo	19.38	–
Wu Wei Zi	19.41	–
Zhi Gan Cao	19.83	–
Dan Dou Chi	19.92	–
Ban Lan Gen	19.96	–
Ce Bai Ye	20.21	–
Pu Gong Ying	20.77	–
Chuang Mu Xiang	21.63	–
Hong Hua	21.63	–
Pi Pa Ye	21.84	–
Che Qian Zi	22.38	–
Chuan Xiong	22.70	–
Bai Xian Pi	22.97	–
Ren Dong Teng	24.03	–
Bai Shao Yao	24.12	–
Tian Hua Fen	24.81	–
Gu Sui Bu	26.18	–
Ze Lan	26.26	–
Lian Zi	27.39	–
Gua Lou	27.81	–
Suan Zao Ren	28.23	–
Ji Li	28.26	–
Lian Qiao	28.30	–
Ren Shen	28.32	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
She Gan	28.48	–
Di Gu Pi	28.60	–
Ji Xue Teng	28.97	–
Ling Zhi	29.02	–
Lu Gen	29.12	–
Gou Qi Zi	29.19	–
Mang Xiao	29.44	–
Mu Zei	30.03	–
Cang Er Zi	30.19	–
Cang Zhu	30.35	–
He Huan Pi	30.53	–
Wu Mei	30.55	–
Zhe Bei Mu	30.67	–
Xi Xian Cao	31.31	–
Guang Huo Xiang	31.84	–
(Fen) Bi Xie	32.13	–
Shen Qu	33.70	–
Shan Yu Rou	33.82	–
Zhu Ling	33.86	–
Yuan Zhi	34.17	–
Huang Qin	34.33	–
Nü Zhen Zi	35.39	–
Tao Ren	36.50	–
Ku Shen	37.10	–
Sheng Jiang	37.55	–
Ban Zhi Lian	37.82	–
Chi Shao (Yao)	37.83	–
Xie Bai	38.17	–
Rou Gui	38.29	–
Jie Geng	38.62	–
Zhi Shi	40.24	–
Gou Teng	40.49	–
Chen Pi	40.71	–
Dang Gui	40.95	–
Ze Xie	42.13	–
Yin Chen Hao	42.30	–
Tai Zi Shen	43.72	–
Fu Ling	43.85	–
Yu Zhu	45.32	–
Jiang Huang	45.36	–
Bai Zi Ren	45.41	–
Xin Yi	45.51	–
Huo Ma Ren	45.71	–
San Qi	45.80	–
Ma Huang Gen	46.27	–
Yi Mu Cao	46.31	–
(Shi) Chang Pu	48.49	–
Niu Bang Zi	48.60	–
Gui Zhi	48.66	–
Zi Su Zi	49.41	–
E Zhu	49.53	–
Dang Gui Wei	49.78	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Bai* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62697	62697	0.00	8.61
62698	62698	0.00	8.01
62837	62837	0.00	7.29
62838	62838	0.00	7.03

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Huang Lian**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60219-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Huang Lian; Coptidis rhizoma

Special notes

When selecting the *Huang Lian* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Huang Lian	3	0	3

Second-stage model

For differentiation of the substance/substance group *Huang Lian* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Huang Lian*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Lian	G079HS261RM1	62287	40	from supplier
PhytoComm	Huang Lian	G079HS261RM1	62288	40	from supplier
PhytoComm	Huang Lian	G079HS261SK1	62707	40	from supplier
PhytoComm	Huang Lian	G079HS261SK1	62708	40	from supplier
PhytoComm	Huang Lian	G079HS261TK2	62919	40	from supplier
PhytoComm	Huang Lian	G079HS261TK2	62920	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Huang Lian*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Huang Lian*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Huang Lian	G079HS261RM1	62287†	20
PhytoComm	Huang Lian	G079HS261RM1	62288†	20
PhytoComm	Huang Lian	G079HS261SK1	62707†	20
PhytoComm	Huang Lian	G079HS261SK1	62708†	20
PhytoComm	Huang Lian	G079HS261TK2	62919†	20
PhytoComm	Huang Lian	G079HS261TK2	62920†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 12 spectra from 5 *Apo-Ident* customers from 4 batches from the substance/substance group *Huang Lian*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Huang Lian	g079h1205321	1
Phytocomm	Huang Lian	G079H1205321	2
PhytoComm	Huang Lian	G079H1205321	4
Phytocomm	Huang Lian	G079H1205421	1
Phytocomm	Huang Lian	G079H1205521	2
PhytoComm	Huang Lian	G079H1205521	2

- 845 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Huang Lian* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Huang Lian* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	0	12	845

The substance/substance group *Huang Lian* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8248 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Huang Lian* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Jiang Cao	12.97	–
Ze Lan	13.13	–
Huang Bai	13.26	–
Hu Zhang	13.61	–
Pu Gong Ying	13.67	–
Sang Ye	14.42	–
Yin Chen Hao	17.38	–
Sha Ren	17.39	–
Tu Fu Ling	20.04	–
Zhi Shi	20.09	–
Han Lian Cao	20.94	–
Bo He	21.10	–
Xin Yi	21.49	–
Shan Yao	21.59	–
Wu Mei	21.62	–
Ma Huang	21.78	–
Bai Hua She She Cao	23.83	–
Yu Xing Cao	24.45	–
Xi Xian Cao	25.14	–
Zhi Ke	25.26	–
Qiang Huo	25.53	–
Fu Zi	25.68	–
Nü Zhen Zi	25.77	–
Jing Jie	25.86	–
Mang Xiao	26.84	–
Qing Pi	26.90	–
Dan Shen	27.26	–
Chai Hu	28.39	–
Hou Po	28.61	–
Fu Pen Zi	29.48	–
Du Zhong	29.63	–
Yin Yang Huo	29.67	–
Yu Jin	30.54	–
Ye Jiao Teng	30.80	–
Hong Jing Tian	31.29	–
Jiao Gu Lan	31.36	–
Xia Ku Cao	31.65	–
Mao Dong Qing	32.47	–
Qing Hao	32.69	–
Yi Mu Cao	33.68	–
Hong Hua	33.79	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Wei Zi	33.95	–
Che Qian Zi	34.48	–
Dan Dou Chi	34.69	–
Zi Hua Di Ding	34.73	–
Rou Cong Rong	36.37	–
Gan Cao	36.38	–
Jiang Huang	37.90	–
Zhi Gan Cao	38.45	–
Jin Yin Hua	39.23	–
Xiang Fu	40.70	–
Pi Pa Ye	40.87	–
Suan Zao Ren	41.10	–
Bai Xian Pi	41.23	–
Chuang Mu Xiang	41.78	–
Huang Qin	41.80	–
Cang Er Zi	42.97	–
Bai Shao Yao	44.26	–
He Huan Pi	44.41	–
Yan Hu Suo	44.99	–
Gou Teng	46.17	–
Sang Ji Shend	46.44	–
Gu Sui Bu	46.76	–
Lian Qiao	47.65	–
Lian Zi	48.74	–
Ren Dong Teng	48.78	–
Chuan Xiong	48.79	–
Ce Bai Ye	49.25	–
Di Gu Pi	49.37	–
Lu Gen	49.88	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Lian* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62287	62287	0.00	13.13
62288	62288	0.00	12.97
62707	62707	0.00	13.86
62708	62708	0.00	13.61
62919	62919	0.00	17.50
62920	62920	0.00	17.38

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all

substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Huang Qi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60154-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Huang Qi; Astragali membranacei radix

Special notes

When selecting the *Huang Qi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Huang Qi	3	0	5

Second-stage model

For differentiation of the substance/substance group *Huang Qi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Huang Qi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Qi	G040H1202822	62383	40	from supplier
PhytoComm	Huang Qi	G040H1202822	62384	40	from supplier
PhytoComm	Huang Qi	G040H1202823	62511	40	from supplier
PhytoComm	Huang Qi	G040H1202823	62512	40	from supplier
PhytoComm	Huang Qi	G040H1202924	62973	40	from supplier
PhytoComm	Huang Qi	G040H1202924	62974	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Huang Qi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Huang Qi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Huang Qi	G040H1202822	62383†	20
PhytoComm	Huang Qi	G040H1202822	62384†	20
PhytoComm	Huang Qi	G040H1202823	62511†	20
PhytoComm	Huang Qi	G040H1202823	62512†	20
PhytoComm	Huang Qi	G040H1202924	62973†	20
PhytoComm	Huang Qi	G040H1202924	62974†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 10 spectra from 5 *Apo-Ident* customers from 6 batches from the substance/substance group *Huang Qi*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Caelo	Huang Qi	g040h1202222	1
Phytocomm	Huang Qi	g040h1202321	2
Phytocomm	Huang Qi	G040H1202321	1
PhytoComm	Huang Qi	G040H1202321	1
Phytocomm	Huang Qi	G040H1202423	2
Phytocomm	Huang Qi	G040H1202521	2
PhytoComm	Huang Qi	G040H1202621	1

- 847 spectra from 13 *Apo-Ident* customers from a total of 513 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Huang Qi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Huang Qi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	7	3	7	840

The substance/substance group *Huang Qi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	99.4785 % (> 98.8910 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Huang Qi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mu Gua	6.43	–
Zhi Gan Cao	8.03	–
Chi Shao (Yao)	10.43	–
Bai Zhu	12.21	–
Jie Geng	15.23	–
(Huai) Niu Xi	18.11	–
Sang Zhi	20.43	–
Long Dan (Cao)	20.69	–
Yuan Zhi	22.26	–
Chuan Lian Zi	22.74	–
Di Gu Pi	22.75	–
Bing Lang	25.88	–
Mang Xiao	26.99	–
Qin Jiao	27.47	–
Tian Hua Fen	28.21	–
Gua Lou	30.50	–
Bai He	31.07	–
Lian Qiao	32.17	–
Gan Cao	32.76	–
Dang Gui	34.16	–
Zhi Ke	34.62	–
Zi Su Zi	34.69	–
Chuan Niu Xi	35.42	–
Bai Zhi	35.53	–
Chuan Mu Tong	35.83	–
Tu Fu Ling	37.45	–
Shan Yao	37.78	–
Ban Zhi Lian	37.99	–
Bai Shao Yao	38.78	–
Lai Fu Zi	39.20	–
Sha Shen (Bei)	39.56	–
(Shi) Chang Pu	40.02	–
Chuan Xiong	40.20	–
Lian Zi	41.03	–
Cang Zhu	41.47	–
E Zhu	41.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Xian Pi	41.86	–
Gou Qi Zi	41.90	–
Gan Jiang	41.96	–
Du Huo	42.57	–
Ji Li	42.63	–
Suan Zao Ren	42.88	–
Ju Hua	43.60	–
Ma Huang	43.60	–
Chai Hu	44.25	–
Huang Qin	44.65	–
She Gan	44.72	–
Zhe Bei Mu	45.09	–
Jin Yin Hua	45.16	–
Ku Shen	45.62	–
Xiang Fu	45.73	–
Rou Gui	45.93	–
Gu Sui Bu	46.02	–
Chen Pi	46.24	–
Yan Hu Suo	46.75	–
Shan Yu Rou	46.78	–
Fu Zi	46.99	–
Ji Xue Teng	47.07	–
Ye Jiao Teng	47.18	–
Fo Shou	47.53	–
Jiao Gu Lan	47.72	–
Huang Bai	47.78	–
Niu Bang Zi	48.15	–
Cang Er Zi	48.22	–
Gou Teng	48.24	–
San Qi	48.86	–
Mai Ya	49.18	–
Ce Bai Ye	49.21	–
Wu Wei Zi	49.28	–
Fu Ling	49.61	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Qi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62383	62383	0.00	7.25
62384	62384	0.00	6.43
62511	62511	0.00	7.80
62512	62512	0.00	8.07
62973	62973	0.00	9.24
62974	62974	0.00	10.20

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Huang Qin**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60012-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Huang Qin; Scutellariae baicalensis radix

Special notes

When selecting the *Huang Qin* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Huang Qin	3	0	5

Second-stage model

For differentiation of the substance/substance group *Huang Qin* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Huang Qin*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Qin	G225HS259RM1	62281	40	from supplier
PhytoComm	Huang Qin	G225HS259RM1	62282	40	from supplier
PhytoComm	Huang Qin	G225HS259SM1	62709	40	from supplier
PhytoComm	Huang Qin	G225HS259SM1	62710	40	from supplier
PhytoComm	Huang Qin	G225HS259TH2	62789	40	from supplier
PhytoComm	Huang Qin	G225HS259TH2	62790	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Huang Qin*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Huang Qin*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Huang Qin	G225HS259RM1	62281†	20
PhytoComm	Huang Qin	G225HS259RM1	62282†	20
PhytoComm	Huang Qin	G225HS259SM1	62709†	20
PhytoComm	Huang Qin	G225HS259SM1	62710†	20
PhytoComm	Huang Qin	G225HS259TH2	62789†	20
PhytoComm	Huang Qin	G225HS259TH2	62790†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 12 spectra from 6 *Apo-Ident* customers from 5 batches from the substance/substance group *Huang Qin*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Herbasin	Huang Qin	g225h120442	1
Phytocomm	Huang Qin	g225h1204121	1
Phytocomm	Huang Qin	G225H1204221	1
PhytoComm	Huang Qin	G225H1204221	1
PhytoComm	Huang Qin	G225H1204322	5
Phytocomm	Huang Qin	G225H1204522	3

- 845 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Huang Qin* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Huang Qin* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	0	12	845

The substance/substance group *Huang Qin* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8248 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Huang Qin* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	23.60	–
Ma Huang	23.99	–
Dang Gui	26.56	–
Jin Yin Hua	26.86	–
Yu Zhu	27.15	–
Lian Qiao	28.46	–
Zhi Ke	28.87	–
Ye Jiao Teng	29.98	–
Jiao Gu Lan	30.28	–
Fu Pen Zi	31.26	–
Hou Po	32.22	–
Gou Qi Zi	32.60	–
Fu Zi	33.23	–
Yin Yang Huo	34.02	–
Tian Hua Fen	34.06	–
Gan Cao	34.28	–
Yan Hu Suo	34.59	–
Tai Zi Shen	34.88	–
Cang Zhu	34.94	–
Bai Shao Yao	36.66	–
Ren Shen	36.91	–
He Huan Pi	37.83	–
Zi Hua Di Ding	37.98	–
Pi Pa Ye	39.96	–
Ban Lan Gen	40.06	–
San Qi	40.21	–
Dang Gui Wei	40.34	–
Guang Huo Xiang	40.49	–
Shan Yu Rou	40.61	–
Lu Gen	40.94	–
Ku Shen	41.61	–
Chen Pi	42.89	–
Chuang Mu Xiang	43.18	–
Hong Hua	43.27	–
Hong Jing Tian	43.31	–
Chi Shao (Yao)	43.89	–
Huang Lian	44.11	–
Wu Wei Zi	44.12	–
Tao Ren	44.30	–
Shan Yao	44.84	–
Rou Gui	45.67	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qiang Huo	46.22	–
Zhe Bei Mu	46.44	–
Di Gu Pi	46.48	–
Gua Lou	46.59	–
Dan Dou Chi	47.34	–
Gou Teng	48.22	–
Ze Xie	48.57	–
Suan Zao Ren	48.64	–
Ci Wu Jia	48.90	–
Ji Xue Teng	49.39	–
Sheng Jiang	49.81	–
Mu Zei	49.84	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Qin* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62281	62281	0.00	24.52
62282	62282	0.00	23.99
62709	62709	0.00	23.72
62710	62710	0.00	23.63
62789	62789	0.00	23.67
62790	62790	0.00	23.60

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Huo Ma Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60234-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Huo Ma Ren; Cannabis semen

Special notes

When selecting the *Huo Ma Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Huo Ma Ren	1	0	1

Second-stage model

For differentiation of the substance/substance group *Huo Ma Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Huo Ma Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huo Ma Ren	G141HS061SP1	62711	40	from supplier
PhytoComm	Huo Ma Ren	G141HS061SP1	62712	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Huo Ma Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Huo Ma Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Huo Ma Ren	G141HS061SP1	62711†	20
PhytoComm	Huo Ma Ren	G141HS061SP1	62712†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Huo Ma Ren*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Huo Ma Ren	G141HS061	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Huo Ma Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Huo Ma Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	2	76	4	24 518
Type B	3	37	3	12 434
Type C	0	0	2	855

The substance/substance group *Huo Ma Ren* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9911 % (> 99.9761 %)	95.0000 % (> 91.2500 %)
Type B	99.9893 % (> 99.9596 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Huo Ma Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
He Huan Pi	3.65	–
Mu Zei	4.76	–
Bai Shao Yao	6.82	–
Gu Sui Bu	6.88	–
Gou Teng	7.34	–
Shen Qu	7.72	–
Tao Ren	7.73	–
Lu Gen	8.23	–
Lian Zi	8.75	–
Gui Zhi	8.99	–
Sheng Jiang	9.05	–
Tai Zi Shen	9.63	–
Ling Zhi	9.69	–
Fu Ling	9.78	–
Ji Xue Teng	9.80	–
Ji Li	9.86	–
Tian Hua Fen	10.49	–
Ci Wu Jia	10.67	–
Zhe Bei Mu	10.69	–
Rou Gui	10.73	–
Bai Xian Pi	11.07	–
Ye Jiao Teng	11.46	–
(Fen) Bi Xie	11.53	–
Ren Dong Teng	11.57	–
She Gan	11.66	–
Bai Zi Ren	11.95	–
Zhu Ling	12.06	–
Yi Yi Ren	12.49	–
Lai Fu Zi	12.50	–
Suan Zao Ren	12.50	–
Fu Xiao Mai	13.24	–
Di Gu Pi	13.35	–
Ban Xia (Jiang)	14.22	–
Ma Huang Gen	14.34	–
Fo Shou	15.47	–
Chuan Xiong	15.48	–
Ce Bai Ye	15.52	–
Yu Jin	15.97	–
Lian Qiao	16.04	–
Ze Xie	16.17	–
Tu Fu Ling	16.20	–
Gua Lou	16.50	–
Pi Pa Ye	16.64	–
Zhu Ru	16.66	–
Shan Yao	16.83	–
Dan Dou Chi	17.80	–
Yan Hu Suo	18.97	–
Dan Shen	19.35	–
Fu Zi	19.55	–
Chen Pi	20.79	–
Fu Pen Zi	21.07	–
Jin Yin Hua	21.40	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chai Hu	21.51	–
Yin Yang Huo	21.67	–
Yuan Zhi	21.71	–
Mao Dong Qing	22.37	–
Guang Huo Xiang	23.48	–
Jiao Gu Lan	23.82	–
Hong Jing Tian	24.48	–
Hou Po	24.82	–
Ban Lan Gen	25.45	–
Yu Zhu	25.90	–
Ma Huang	26.09	–
Fu Shen	26.29	–
Che Qian Zi	26.37	–
Cang Zhu	26.37	–
Dang Gui Wei	27.29	–
Gan Cao	27.42	–
Zhi Ke	28.58	–
Qiang Huo	29.09	–
Jie Geng	29.18	–
Ren Shen	29.24	–
Dang Gui	30.18	–
San Qi	30.82	–
Jing Jie	32.36	–
Mang Xiao	32.98	–
Zi Hua Di Ding	34.76	–
Wu Wei Zi	34.85	–
Huang Qin	34.91	–
Zhi Mu	35.28	–
Shan Yu Rou	35.94	–
Sha Ren	36.78	–
Huang Lian	37.16	–
Chuang Mu Xiang	38.91	–
Sang Zhi	40.04	–
Gou Qi Zi	40.11	–
Mu Gua	40.28	–
Qing Pi	40.30	–
Long Yan Rou	40.31	–
Zhi Gan Cao	40.45	–
Du Zhong	40.61	–
Bo He	42.29	–
Ku Shen	42.78	–
Xie Bai	44.32	–
Huang Bai	46.62	–
Chi Shao (Yao)	47.45	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huo Ma Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62711	62711	0.00	4.91
62712	62712	0.00	3.65

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ji Li**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60110-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ji Li; Tribuli terrestris fructus

Special notes

When selecting the *Ji Li* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ji Li	3	0	1

Second-stage model

For differentiation of the substance/substance group *Ji Li* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ji Li*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ji Li	G240H1410821	62319	40	from supplier
PhytoComm	Ji Li	G240H1410821	62320	40	from supplier
PhytoComm	Ji Li	G240HS305SK1	62545	40	from supplier
PhytoComm	Ji Li	G240HS305SK1	62546	40	from supplier
PhytoComm	Ji Li	G240HS305SQ1	62721	40	from supplier
PhytoComm	Ji Li	G240HS305SQ1	62722	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Ji Li*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Ji Li*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ji Li	G240H1410821	62319†	20
PhytoComm	Ji Li	G240H1410821	62320†	20
PhytoComm	Ji Li	G240HS305SK1	62545†	20
PhytoComm	Ji Li	G240HS305SK1	62546†	20
PhytoComm	Ji Li	G240HS305SQ1	62721†	20
PhytoComm	Ji Li	G240HS305SQ1	62722†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Ji Li*.

- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ji Li	G240H1410421	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ji Li* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ji Li* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	240	0	24 359
Type B	3	114	6	12 354
Type C	0	0	1	856

The substance/substance group *Ji Li* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9964 % (> 99.9814 %)	100.0000 % (> 97.5000 %)
Type B	99.9786 % (> 99.9487 %)	95.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ji Li* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lai Fu Zi	4.27	–
Ling Zhi	4.59	–
Lian Qiao	5.21	–
Bai Shao Yao	5.96	–
Tian Hua Fen	6.09	–
Shen Qu	6.37	–
Bai Zi Ren	7.02	–
Mu Zei	7.23	–
Sheng Jiang	7.29	–
Zhe Bei Mu	7.29	–
Tai Zi Shen	7.47	–
Lian Zi	7.97	–
Huo Ma Ren	8.17	–
He Huan Pi	8.40	–
Fu Zi	8.78	–
Gua Lou	8.87	–
Pi Pa Ye	9.63	–
Zhu Ru	9.88	–
Ren Dong Teng	10.05	–
Shan Yao	10.07	–
Ci Wu Jia	10.16	–
Gu Sui Bu	10.26	–
Tao Ren	10.48	–
Fu Ling	11.69	–
Di Gu Pi	11.99	–
Gui Zhi	12.07	–
Chuan Xiong	12.43	–
Hou Po	12.61	–
Ye Jiao Teng	12.96	–
Mao Dong Qing	13.02	–
Fu Xiao Mai	13.39	–
Yu Jin	13.39	–
Jiang Huang	13.41	–
Yuan Zhi	13.80	–
Che Qian Zi	14.26	–
Ze Xie	14.27	–
Chai Hu	14.29	–
Suan Zao Ren	14.35	–
Gou Teng	14.63	–
Ban Xia (Jiang)	14.78	–
Fu Pen Zi	15.03	–
E Zhu	15.10	–
Yi Yi Ren	15.46	–
(Bai) Dou Kou	15.51	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Xian Pi	15.71	–
Dan Shen	15.76	–
Jiao Gu Lan	15.85	–
Tu Fu Ling	16.23	–
Lu Gen	16.27	–
Rou Gui	16.30	–
Dan Dou Chi	16.35	–
Ji Xue Teng	16.38	–
Chen Pi	16.66	–
Ce Bai Ye	16.70	–
Cang Zhu	16.92	–
Sha Ren	17.40	–
Zhu Ling	17.42	–
Fo Shou	17.77	–
Dang Gui Wei	17.89	–
She Gan	18.07	–
Jin Yin Hua	18.09	–
San Qi	18.65	–
Yan Hu Suo	18.94	–
Guang Huo Xiang	19.65	–
Ban Lan Gen	20.46	–
Ren Shen	20.82	–
(Fen) Bi Xie	20.91	–
Jie Geng	21.05	–
Dang Gui	21.11	–
Yu Zhu	21.24	–
Chuan Lian Zi	21.86	–
Zi Su Zi	21.91	–
Yin Yang Huo	22.36	–
Du Huo	22.91	–
Sha Shen (Bei)	22.94	–
Ma Huang Gen	23.18	–
Qiang Huo	23.33	–
Zhi Mu	23.49	–
Wu Wei Zi	23.66	–
Zhi Gan Cao	24.75	–
Ku Shen	24.90	–
Chuan Niu Xi	25.00	–
Hong Jing Tian	25.25	–
Mang Xiao	25.40	–
Xie Bai	25.55	–
Gan Cao	25.69	–
Gan Jiang	26.07	–
Niu Bang Zi	26.77	–
Shan Yu Rou	27.00	–
Gou Qi Zi	27.14	–
Zhi Ke	27.39	–
Qing Pi	27.59	–
Sang Zhi	27.60	–
Cang Er Zi	28.50	–
Mu Gua	28.54	–
Huang Qin	28.72	–
(Shi) Chang Pu	28.81	–
Bai Hua She She Cao	29.08	–
(Huai) Niu Xi	29.18	–
Wu Yao	29.53	–
Ma Huang	29.91	–
Xiang Fu	30.18	–
Du Zhong	30.20	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mi Huan Jun	30.30	–
Zi Hua Di Ding	31.42	–
Jing Jie	32.04	–
Chuang Mu Xiang	32.38	–
Fu Shen	32.95	–
Long Yan Rou	33.39	–
Bo He	33.61	–
Huang Lian	34.20	–
Xiao Hui Xiang	35.52	–
Bai Zhi	35.65	–
Huang Qi	37.07	–
Huang Bai	38.13	–
Ba Ji Tian	38.28	–
Mu Dan Pi	38.56	–
Ju Hua	39.61	–
Chi Shao (Yao)	39.84	–
Sang Ji Shend	40.61	–
Xin Yi	40.70	–
Chuan Mu Tong	42.55	–
Xu Duan	42.82	–
Da Zao	43.07	–
Qin Jiao	43.15	–
Wang Bu Liu Xing	43.22	–
Yi Mu Cao	43.60	–
Fang Feng	45.37	–
Mai Ya	46.27	–
Xi Xian Cao	46.37	–
He Shou Wu	46.76	–
Ze Lan	47.20	–
Hong Hua	47.21	–
Long Dan (Cao)	47.83	–
Qing Hao	48.08	–
Bai Zhu	48.69	–
Bai Jiang Cao	49.17	–
Pu Gong Ying	49.37	–
Yin Chen Hao	50.48	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ji Li* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62319	62319	0.00	13.08
62320	62320	0.00	13.22
62545	62545	0.00	4.83
62546	62546	0.00	4.59
62721	62721	0.00	4.27
62722	62722	0.00	4.28

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ji Xue Teng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60099-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ji Xue Teng; Spatholobi caulis

Special notes

When selecting the *Ji Xue Teng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ji Xue Teng	2	0	2

Second-stage model

For differentiation of the substance/substance group *Ji Xue Teng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ji Xue Teng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ji Xue Teng	G159HS337SK1	62559	40	from supplier
PhytoComm	Ji Xue Teng	G159HS337SK1	62560	40	from supplier
PhytoComm	Ji Xue Teng	G159HS337SQ1	62723	40	from supplier
PhytoComm	Ji Xue Teng	G159HS337SQ1	62724	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ji Xue Teng*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ji Xue Teng*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ji Xue Teng	G159HS337SK1	62559†	20
PhytoComm	Ji Xue Teng	G159HS337SK1	62560†	20
PhytoComm	Ji Xue Teng	G159HS337SQ1	62723†	20
PhytoComm	Ji Xue Teng	G159HS337SQ1	62724†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Ji Xue Teng*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ji Xue Teng	g159h2101122	1
Phytocomm	Ji Xue Teng	G159H2101423	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ji Xue Teng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ji Xue Teng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	154	6	24 437
Type B	3	78	2	12 394
Type C	0	0	4	853

The substance/substance group *Ji Xue Teng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9857 % (> 99.9707 %)	96.2500 % (> 94.3750 %)
Type B	99.9714 % (> 99.9416 %)	97.5000 % (> 93.7500 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ji Xue Teng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lu Gen	3.90	–
Ye Jiao Teng	4.17	–
He Huan Pi	5.43	–
Rou Gui	6.12	–
Ma Huang Gen	6.86	–
Ren Dong Teng	7.91	–
Gou Teng	9.40	–
(Fen) Bi Xie	10.23	–
Mu Zei	10.46	–
Fo Shou	10.76	–
Gui Zhi	10.82	–
Bai Shao Yao	11.47	–
Tai Zi Shen	11.59	–
She Gan	11.73	–
Huo Ma Ren	11.74	–
Fu Ling	12.06	–
Bai Xian Pi	12.12	–
Dan Dou Chi	14.15	–
Yi Yi Ren	14.19	–
Sheng Jiang	14.47	–
Gu Sui Bu	15.20	–
Tu Fu Ling	15.85	–
Ling Zhi	16.41	–
Tao Ren	16.51	–
Shen Qu	16.64	–
Dan Shen	16.74	–
Lian Zi	16.88	–
Ci Wu Jia	17.10	–
Zhu Ling	17.11	–
Ban Xia (Jiang)	17.36	–
Fu Shen	17.64	–
Suan Zao Ren	18.02	–
Ma Huang	19.07	–
Yin Yang Huo	19.62	–
Yan Hu Suo	19.96	–
Guang Huo Xiang	20.37	–
Yu Jin	20.39	–
Fu Xiao Mai	20.69	–
Lai Fu Zi	20.70	–
Zhe Bei Mu	20.98	–
Di Gu Pi	21.79	–
Chuan Xiong	22.18	–
Ji Li	22.77	–
Lian Qiao	23.25	–
Ce Bai Ye	23.32	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	23.57	–
Tian Hua Fen	23.58	–
Bai Zi Ren	24.59	–
Ze Xie	25.45	–
Pi Pa Ye	25.73	–
Chai Hu	25.90	–
Fu Zi	25.94	–
Gua Lou	26.00	–
Mao Dong Qing	26.13	–
Hong Jing Tian	26.55	–
Fu Pen Zi	26.70	–
Chen Pi	27.05	–
Jing Jie	27.85	–
Gan Cao	27.91	–
Zhu Ru	27.93	–
Zhi Ke	28.32	–
Shan Yao	28.56	–
Jiao Gu Lan	28.70	–
Hou Po	28.73	–
Yu Zhu	29.54	–
Qiang Huo	30.39	–
Ban Lan Gen	30.41	–
Mang Xiao	30.60	–
Yuan Zhi	30.73	–
Cang Zhu	31.62	–
Zi Hua Di Ding	31.66	–
Sha Ren	32.41	–
Ren Shen	32.88	–
Che Qian Zi	32.98	–
Huang Lian	34.02	–
Huang Qin	34.53	–
Dang Gui Wei	35.06	–
Jie Geng	35.13	–
Dang Gui	35.16	–
Wu Wei Zi	39.22	–
Zhi Mu	39.30	–
Du Zhong	39.33	–
Long Yan Rou	39.67	–
Bo He	39.87	–
Sang Zhi	39.93	–
Chuang Mu Xiang	41.63	–
Gou Qi Zi	42.40	–
San Qi	42.64	–
Shan Yu Rou	42.83	–
Huang Bai	42.96	–
Chi Shao (Yao)	44.21	–
Zhi Gan Cao	44.51	–
Mu Gua	46.35	–
Ban Zhi Lian	47.41	–
Qing Pi	47.73	–
Hong Hua	49.23	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ji Xue Teng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62559	62559	0.00	4.45
62560	62560	0.00	4.17
62723	62723	0.00	3.90
62724	62724	0.00	4.96

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jiang Huang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	`60181-2022-02-15`
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jiang Huang; Curcumae longae rhizoma

Special notes

When selecting the *Jiang Huang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jiang Huang	1	0	2

Second-stage model

For differentiation of the substance/substance group *Jiang Huang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jiang Huang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jiang Huang	G085H1716821	62519	40	from supplier
PhytoComm	Jiang Huang	G085H1716821	62520	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Jiang Huang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Jiang Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jiang Huang	G085H1716821	62519†	20
PhytoComm	Jiang Huang	G085H1716821	62520†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Jiang Huang*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Jiang Huang	G085H1716321	1
Phytocomm	Jiang Huang	G085H1716422	3
PhytoComm	Jiang Huang	G085H1716422	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jiang Huang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jiang Huang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	3	0	5	849

The substance/substance group *Jiang Huang* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.3023 % (> 98.7155 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jiang Huang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sha Ren	14.08	–
(Bai) Dou Kou	19.18	–
Ji Li	19.33	–
He Huan Pi	20.50	–
Bai Hua She She Cao	20.52	–
E Zhu	23.61	–
Yan Hu Suo	23.77	–
Shan Yao	24.30	–
Wang Bu Liu Xing	25.58	–
Yi Mu Cao	25.83	–
Cang Er Zi	27.87	–
Ze Lan	28.11	–
Niu Bang Zi	28.18	–
Chen Pi	28.68	–
Huang Lian	29.64	–
Mang Xiao	29.81	–
Xin Yi	30.35	–
Xiang Fu	31.51	–
Hu Zhang	32.56	–
Zhi Ke	32.86	–
Yin Chen Hao	33.95	–
Zi Su Zi	34.07	–
Wu Yao	34.66	–
Du Zhong	35.03	–
Di Gu Pi	35.15	–
Mao Dong Qing	35.51	–
Xi Xian Cao	35.75	–
Pu Gong Ying	35.80	–
Sang Ye	36.49	–
(Shi) Chang Pu	36.59	–
Suan Zao Ren	37.42	–
Xiao Hui Xiang	37.61	–
Gan Jiang	38.08	–
Gu Sui Bu	38.19	–
Qiang Huo	40.77	–
Sang Ji Shend	40.90	–
Gua Lou	42.04	–
Nü Zhen Zi	42.10	–
Bo He	42.67	–
Sha Shen (Bei)	45.05	–
Hong Jing Tian	45.66	–
Dan Dou Chi	45.66	–
Mi Huan Jun	45.74	–
Tian Hua Fen	46.11	–
Tu Fu Ling	46.24	–
Jiao Gu Lan	46.28	–
Xia Ku Cao	46.45	–
Shan Yu Rou	46.60	–
Xu Duan	46.64	–
Han Lian Cao	47.02	–
He Shou Wu	47.38	–
Yu Xing Cao	47.95	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Qiao	48.10	–
Chuan Niu Xi	48.24	–
Dang Gui	48.46	–
Wu Wei Zi	49.21	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jiang Huang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62519	62519	0.00	14.08
62520	62520	0.00	14.52

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jiao Gu Lan**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60332-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jiao Gu Lan; Gynostemma herba

Special notes

When selecting the *Jiao Gu Lan* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jiao Gu Lan	3	0	1

Second-stage model

For differentiation of the substance/substance group *Jiao Gu Lan* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jiao Gu Lan*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jiao Gu Lan	G065HS424RR1	62441	40	from supplier
PhytoComm	Jiao Gu Lan	G065HS424RR1	62442	40	from supplier
PhytoComm	Jiao Gu Lan	G065HS424SG1	62665	40	from supplier
PhytoComm	Jiao Gu Lan	G065HS424SG1	62666	40	from supplier
PhytoComm	Jiao Gu Lan	G065HS424TL1	62907	40	from supplier
PhytoComm	Jiao Gu Lan	G065HS424TL1	62908	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Jiao Gu Lan*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Jiao Gu Lan*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jiao Gu Lan	G065HS424RR1	62441†	20
PhytoComm	Jiao Gu Lan	G065HS424RR1	62442†	20
PhytoComm	Jiao Gu Lan	G065HS424SG1	62665†	20
PhytoComm	Jiao Gu Lan	G065HS424SG1	62666†	20
PhytoComm	Jiao Gu Lan	G065HS424TL1	62907†	20
PhytoComm	Jiao Gu Lan	G065HS424TL1	62908†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Jiao Gu Lan*.

- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Jiao Gu Lan	G065FB1276421	3

- 854 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jiao Gu Lan* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jiao Gu Lan* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	0	3	854

The substance/substance group *Jiao Gu Lan* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jiao Gu Lan* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	7.74	–
Shan Yao	9.71	–
Cang Zhu	10.82	–
Zhi Gan Cao	10.94	–
Suan Zao Ren	11.07	–
Lian Qiao	11.71	–
Dan Shen	11.94	–
Zi Hua Di Ding	12.09	–
Hong Hua	13.80	–
Ji Li	14.01	–
Chai Hu	14.22	–
Wu Wei Zi	14.26	–
Xie Bai	14.46	–
Ku Shen	14.80	–
Hong Jing Tian	14.88	–
Qing Pi	15.04	–
Zhi Ke	15.77	–
Mao Dong Qing	15.88	–
Chuan Xiong	16.07	–
Ling Zhi	16.42	–
Qing Hao	16.45	–
Ban Lan Gen	16.61	–
Gou Qi Zi	17.18	–
Bai Zi Ren	18.02	–
Shan Yu Rou	18.04	–
Guang Huo Xiang	18.51	–
Gua Lou	18.54	–
Dan Dou Chi	18.55	–
Gan Cao	18.63	–
Chuang Mu Xiang	18.86	–
Zhe Bei Mu	18.95	–
Tian Hua Fen	19.03	–
Pi Pa Ye	19.31	–
Chen Pi	19.40	–
Yuan Zhi	20.12	–
Hou Po	20.93	–
Ren Shen	21.38	–
Lian Zi	21.74	–
Qiang Huo	22.20	–
She Gan	22.45	–
Cang Er Zi	22.58	–
Yan Hu Suo	22.59	–
Ren Dong Teng	22.77	–
Bo He	22.91	–
Gu Sui Bu	23.44	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Shao Yao	24.46	–
Huang Bai	24.73	–
Fu Zi	25.01	–
Sang Ye	25.31	–
Du Zhong	25.31	–
Ye Jiao Teng	25.41	–
Fu Ling	25.53	–
Yin Yang Huo	25.63	–
Shen Qu	25.98	–
Ma Huang	26.12	–
San Qi	26.21	–
Sha Ren	26.30	–
Fu Pen Zi	26.40	–
Mu Zei	26.64	–
Di Gu Pi	26.75	–
Yu Zhu	27.05	–
He Huan Pi	27.45	–
Ze Xie	27.96	–
Jie Geng	28.02	–
Lu Gen	28.24	–
Dang Gui	28.40	–
Ce Bai Ye	28.68	–
Huo Ma Ren	28.85	–
Bai Xian Pi	29.12	–
Tai Zi Shen	29.51	–
Sheng Jiang	29.82	–
Huang Lian	30.60	–
Zhi Mu	30.85	–
Yu Jin	30.85	–
Wu Mei	30.86	–
Che Qian Zi	31.34	–
Chi Shao (Yao)	31.37	–
Jing Jie	31.76	–
Zhu Ru	31.84	–
Mang Xiao	31.93	–
Ji Xue Teng	32.32	–
Zhu Ling	32.85	–
Gou Teng	33.76	–
Bai Jiang Cao	34.24	–
Huang Qin	34.70	–
Mu Gua	35.45	–
Long Yan Rou	35.77	–
Tao Ren	36.63	–
Gui Zhi	37.05	–
Dang Gui Wei	37.15	–
Tu Fu Ling	38.03	–
Sang Zhi	39.48	–
Rou Gui	40.35	–
(Fen) Bi Xie	40.81	–
Lai Fu Zi	40.82	–
Ci Wu Jia	41.76	–
Niu Bang Zi	42.34	–
Yi Yi Ren	43.28	–
Pu Gong Ying	43.33	–
Ze Lan	45.18	–
Xi Xian Cao	46.05	–
Fu Xiao Mai	46.66	–
Ma Huang Gen	46.98	–
Ban Xia (Jiang)	49.84	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Nü Zhen Zi	49.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jiao Gu Lan* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62441	62441	0.00	7.74
62442	62442	0.00	9.20
62665	62665	0.00	9.71
62666	62666	0.00	9.99
62907	62907	0.00	10.21
62908	62908	0.00	9.88

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jie Geng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60034-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jie Geng; Platycodi grandiflori radix

Special notes

When selecting the *Jie Geng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jie Geng	3	0	2

Second-stage model

For differentiation of the substance/substance group *Jie Geng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jie Geng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jie Geng	G193H1015821	62365	40	from supplier
PhytoComm	Jie Geng	G193H1015821	62366	40	from supplier
PhytoComm	Jie Geng	G193H1015921	62843	40	from supplier
PhytoComm	Jie Geng	G193H1015921	62844	40	from supplier
PhytoComm	Jie Geng	G193HS221TP1	63029	40	from supplier
PhytoComm	Jie Geng	G193HS221TP1	63030	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Jie Geng*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Jie Geng*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jie Geng	G193H1015821	62365†	20
PhytoComm	Jie Geng	G193H1015821	62366†	20
PhytoComm	Jie Geng	G193H1015921	62843†	20
PhytoComm	Jie Geng	G193H1015921	62844†	20
PhytoComm	Jie Geng	G193HS221TP1	63029†	20
PhytoComm	Jie Geng	G193HS221TP1	63030†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Jie Geng*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Jie Geng	g193h1015221	1
Phytocomm	Jie Geng	G193H1015521	4

- 852 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jie Geng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jie Geng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	5	120	0	12 352
Type C	4	2	3	848

The substance/substance group *Jie Geng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	99.9786 % (> 99.9487 %)	100.0000 % (> 95.0000 %)
Type C	99.3798 % (> 98.7930 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jie Geng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dang Gui Wei	4.38	–
Gua Lou	7.20	–
Cang Zhu	7.98	–
Dang Gui	8.90	–
Yuan Zhi	9.33	–
Bai Zhu	9.35	–
Zhi Mu	9.39	–
Bai Shao Yao	10.04	–
Chuan Lian Zi	10.46	–
Yu Zhu	10.59	–
Fo Shou	11.27	–
Sang Zhi	11.42	–
Di Gu Pi	12.82	–
Mu Gua	13.29	–
San Qi	13.39	–
Ye Jiao Teng	13.59	–
Zhi Gan Cao	13.60	–
Pi Pa Ye	14.37	–
Long Dan (Cao)	15.01	–
Gou Qi Zi	15.37	–
Tai Zi Shen	15.48	–
Huang Qi	15.84	–
Ren Shen	15.86	–
Chen Pi	16.07	–
Zhe Bei Mu	16.12	–
Chuan Xiong	16.28	–
Tian Hua Fen	16.61	–
Qin Jiao	17.39	–
Mu Zei	17.49	–
Ji Li	17.68	–
Jin Yin Hua	17.69	–
Bai He	17.78	–
Ban Lan Gen	17.81	–
Shen Qu	18.16	–
He Huan Pi	18.20	–
Long Yan Rou	18.33	–
Ren Dong Teng	18.63	–
Shan Yao	18.69	–
Lian Qiao	19.32	–
Ci Wu Jia	19.32	–
Shan Yu Rou	19.77	–
Lai Fu Zi	19.85	–
Ku Shen	20.33	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhu Ru	21.11	–
Rou Gui	21.14	–
Lian Zi	21.43	–
Bai Zi Ren	21.61	–
Ze Xie	21.64	–
Xie Bai	22.08	–
Jiao Gu Lan	22.78	–
Tao Ren	22.86	–
Chi Shao (Yao)	23.07	–
Ling Zhi	23.08	–
Chuang Mu Xiang	24.47	–
Sheng Jiang	24.62	–
Ban Xia (Jiang)	25.23	–
Tu Fu Ling	25.28	–
Zhi Ke	25.30	–
Suan Zao Ren	25.36	–
Gou Teng	25.41	–
Dan Dou Chi	25.60	–
Gan Jiang	25.78	–
Fu Ling	25.99	–
Sha Shen (Bei)	26.21	–
Lu Gen	26.44	–
(Huai) Niu Xi	26.44	–
Zi Su Zi	26.47	–
Fu Xiao Mai	26.64	–
Fu Pen Zi	26.66	–
Huo Ma Ren	26.82	–
Hong Jing Tian	27.22	–
Fu Zi	27.29	–
Gui Zhi	27.67	–
Ma Huang	27.68	–
Huang Qin	28.08	–
Gan Cao	28.11	–
Ji Xue Teng	28.22	–
Chuan Mu Tong	29.22	–
Hou Po	29.35	–
Mang Xiao	29.78	–
Mai Men Dong	30.00	–
Wu Wei Zi	30.27	–
Dan Shen	30.88	–
Bai Xian Pi	31.07	–
Gu Sui Bu	31.21	–
E Zhu	31.81	–
Cang Er Zi	33.59	–
Mao Dong Qing	33.93	–
Yin Yang Huo	34.07	–
Yi Yi Ren	34.23	–
Ban Zhi Lian	34.26	–
Chai Hu	34.76	–
(Shi) Chang Pu	35.07	–
Che Qian Zi	35.12	–
Guang Huo Xiang	35.15	–
Yan Hu Suo	37.24	–
Ma Huang Gen	37.28	–
Bai Zhi	37.49	–
Chuan Niu Xi	37.86	–
Bing Lang	38.19	–
(Fen) Bi Xie	38.50	–
Mai Ya	38.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ce Bai Ye	39.09	–
Mi Huan Jun	39.26	–
Niu Bang Zi	39.37	–
Fu Shen	40.77	–
Yu Jin	41.05	–
Qiang Huo	41.38	–
She Gan	41.57	–
Zi Hua Di Ding	43.51	–
Zhu Ling	44.64	–
Bo He	45.23	–
Jiang Huang	47.64	–
Qing Pi	47.87	–
Xiao Hui Xiang	48.86	–
Da Zao	49.04	–
Hong Hua	49.12	–
Sha Ren	49.13	–
Huang Lian	49.26	–
Dong Gua Zi	49.89	–
Pu Gong Ying	49.98	–
Jing Jie	50.15	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jie Geng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62365	62365	0.00	9.35
62366	62366	0.00	10.46
62843	62843	0.00	12.45
62844	62844	0.00	13.89
63029	63029	0.00	4.38
63030	63030	0.00	5.84

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jin Qian Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50293-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jin Qian Cao; Lysimachiae herba

Special notes

When selecting the *Jin Qian Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jin Qian Cao	1	0	1

Second-stage model

For differentiation of the substance/substance group *Jin Qian Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jin Qian Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jin Qian Cao	G151H0847821	62465	40	from supplier
PhytoComm	Jin Qian Cao	G151H0847821	62466	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Jin Qian Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Jin Qian Cao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jin Qian Cao	G151H0847821	62465†	20
PhytoComm	Jin Qian Cao	G151H0847821	62466†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Jin Qian Cao*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Jin Qian Cao	g151h0847122	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jin Qian Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jin Qian Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	1	0	1	855

The substance/substance group *Jin Qian Cao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8837 % (> 99.3020 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jin Qian Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sang Ye	16.68	–
Ding Xiang	17.17	–
Wu Jia Pi	19.90	–
Jing Jie	20.42	–
Mang Xiao	21.14	–
Du Zhong	24.32	–
Sang Ji Shend	26.23	–
Gou Teng	27.99	–
Yu Xing Cao	31.86	–
Xuan Shen	33.18	–
Hong Jing Tian	33.48	–
Dan Zhu Ye	36.42	–
Ge Gen	36.58	–
Xuan Fu Hua	38.02	–
Yin Chen Hao	41.19	–
Nü Zhen Zi	42.75	–
Xian Mao	42.77	–
Zhi Shi	43.03	–
He Shou Wu	48.59	–
Hua Shi	49.02	–
(Sheng) Di Huang	49.10	–
Yi Mu Cao	52.63	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jin Qian Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62465	62465	0.00	16.68
62466	62466	0.00	17.07

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jin Yin Hua**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60350-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jin Yin Hua; Lonicerae japonicae flos

Special notes

When selecting the *Jin Yin Hua* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jin Yin Hua	2	0	0

Second-stage model

For differentiation of the substance/substance group *Jin Yin Hua* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jin Yin Hua*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jin Yin Hua	G145HS311SK1	62577	40	from supplier
PhytoComm	Jin Yin Hua	G145HS311SK1	62578	40	from supplier
PhytoComm	Jin Yin Hua	G145HS311TH2	62897	40	from supplier
PhytoComm	Jin Yin Hua	G145HS311TH2	62898	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Jin Yin Hua*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Jin Yin Hua*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jin Yin Hua	G145HS311SK1	62577†	20
PhytoComm	Jin Yin Hua	G145HS311SK1	62578†	20
PhytoComm	Jin Yin Hua	G145HS311TH2	62897†	20
PhytoComm	Jin Yin Hua	G145HS311TH2	62898†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Jin Yin Hua*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jin Yin Hua* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jin Yin Hua* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	2	80	0	12 395
Type C	0	0	0	857

The substance/substance group *Jin Yin Hua* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9857 % (> 99.9559 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jin Yin Hua* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ma Huang	8.86	–
Jiao Gu Lan	9.61	–
Ban Lan Gen	10.51	–
Chuan Xiong	11.00	–
Pi Pa Ye	11.35	–
Zhe Bei Mu	11.44	–
Yu Zhu	11.45	–
Lian Qiao	11.94	–
Yan Hu Suo	12.11	–
Ze Xie	12.16	–
Fu Zi	12.18	–
Tian Hua Fen	12.28	–
Gua Lou	12.51	–
Tai Zi Shen	12.52	–
Hou Po	12.56	–
Cang Zhu	12.69	–
Guang Huo Xiang	12.95	–
Dan Dou Chi	13.41	–
Ku Shen	13.59	–
Shan Yao	13.61	–
Gou Qi Zi	13.84	–
Zhi Gan Cao	13.94	–
Zhi Ke	13.96	–
Wu Wei Zi	14.49	–
Dang Gui	14.55	–
Ren Dong Teng	14.59	–
Bai Shao Yao	14.82	–
Ye Jiao Teng	15.03	–
He Huan Pi	16.43	–
Ling Zhi	16.68	–
Chai Hu	16.74	–
Huang Qin	16.75	–
Dang Gui Wei	16.87	–
Ren Shen	17.08	–
Ji Li	17.22	–
Lu Gen	17.68	–
Mu Zei	17.68	–
Suan Zao Ren	17.69	–
Fu Pen Zi	17.77	–
Dan Shen	17.84	–
Qing Pi	17.90	–
Chen Pi	17.95	–
Tao Ren	18.17	–
Zi Hua Di Ding	18.23	–
Hong Jing Tian	18.57	–
Xie Bai	18.91	–
Gan Cao	19.12	–
San Qi	19.80	–
Yuan Zhi	19.94	–
Gui Zhi	20.00	–
Gu Sui Bu	20.29	–
Jie Geng	20.41	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
She Gan	20.46	–
Bai Zi Ren	20.71	–
Chuang Mu Xiang	20.97	–
Di Gu Pi	21.01	–
Shan Yu Rou	21.22	–
Rou Gui	21.25	–
Lian Zi	21.69	–
Ci Wu Jia	22.07	–
Qiang Huo	22.11	–
Mao Dong Qing	22.16	–
Ji Xue Teng	22.21	–
Yin Yang Huo	22.72	–
Lai Fu Zi	23.56	–
Ban Xia (Jiang)	24.03	–
Huo Ma Ren	24.08	–
Zhi Mu	24.10	–
Shen Qu	24.26	–
Ce Bai Ye	24.84	–
Sheng Jiang	25.38	–
Gou Teng	25.51	–
Long Yan Rou	26.09	–
Che Qian Zi	26.15	–
Hong Hua	26.97	–
Tu Fu Ling	26.98	–
Bai Xian Pi	27.50	–
Yu Jin	27.84	–
Bo He	27.93	–
Fo Shou	27.99	–
Fu Ling	28.17	–
Zhu Ling	28.70	–
Mu Gua	29.40	–
Du Zhong	29.57	–
Zhu Ru	29.59	–
Huang Lian	29.97	–
Sha Ren	30.23	–
Mang Xiao	30.52	–
Chi Shao (Yao)	31.77	–
(Fen) Bi Xie	31.99	–
Sang Ye	32.35	–
Yi Yi Ren	32.54	–
Huang Bai	32.59	–
Jing Jie	32.98	–
Cang Er Zi	33.10	–
Fu Xiao Mai	33.21	–
Qing Hao	35.47	–
Fu Shen	37.70	–
Sang Zhi	38.23	–
Ma Huang Gen	39.94	–
Bai Jiang Cao	40.67	–
Ze Lan	49.95	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jin Yin Hua* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62577	62577	0.00	9.88
62578	62578	0.00	9.61
62897	62897	0.00	10.40
62898	62898	0.00	8.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jing Jie**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60088-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jing Jie; Schizonepetae tenuifoliae herba

Special notes

When selecting the *Jing Jie* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jing Jie	3	0	1

Second-stage model

For differentiation of the substance/substance group *Jing Jie* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jing Jie*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jing Jie	G221H1038822	62497	40	from supplier
PhytoComm	Jing Jie	G221H1038822	62498	40	from supplier
PhytoComm	Jing Jie	G221H1038823	62751	40	from supplier
PhytoComm	Jing Jie	G221H1038823	62752	40	from supplier
PhytoComm	Jing Jie	G221HS218TH1	62785	40	from supplier
PhytoComm	Jing Jie	G221HS218TH1	62786	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Jing Jie*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Jing Jie*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jing Jie	G221H1038822	62497†	20
PhytoComm	Jing Jie	G221H1038822	62498†	20
PhytoComm	Jing Jie	G221H1038823	62751†	20
PhytoComm	Jing Jie	G221H1038823	62752†	20
PhytoComm	Jing Jie	G221HS218TH1	62785†	20
PhytoComm	Jing Jie	G221HS218TH1	62786†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Jing Jie*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Jing Jie	g221h1038221	1
Phytocomm	Jing Jie	G221H1038221	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jing Jie* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jing Jie* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	238	2	24 359
Type B	1	117	3	12 356
Type C	4	0	2	851

The substance/substance group *Jing Jie* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9964 % (> 99.9814 %)	99.1667 % (> 97.9167 %)
Type B	99.9857 % (> 99.9558 %)	97.5000 % (> 95.0000 %)
Type C	99.6512 % (> 99.0662 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jing Jie* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hou Po	4.86	–
Mao Dong Qing	5.71	–
Chai Hu	5.85	–
Qiang Huo	5.98	–
Yu Jin	6.03	–
Yin Yang Huo	6.13	–
Du Zhong	6.97	–
Zi Hua Di Ding	7.58	–
Dan Dou Chi	7.84	–
Bai Xian Pi	7.91	–
Fu Zi	7.97	–
Ling Zhi	8.39	–
Ce Bai Ye	9.53	–
Che Qian Zi	9.64	–
Shan Yao	9.89	–
Bo He	10.32	–
Yan Hu Suo	10.39	–
Sha Ren	10.90	–
Shen Qu	11.03	–
Fu Pen Zi	11.72	–
Guang Huo Xiang	12.03	–
Dan Shen	12.06	–
Ye Jiao Teng	12.67	–
Ren Dong Teng	12.80	–
Chuan Xiong	13.62	–
Zhu Ling	13.73	–
Pi Pa Ye	13.74	–
Yin Chen Hao	13.89	–
Qing Pi	14.63	–
Qing Hao	15.03	–
Gu Sui Bu	15.32	–
Hong Jing Tian	15.55	–
Dan Zhu Ye	15.76	–
Ji Li	15.98	–
Lian Zi	16.05	–
Huang Bai	16.16	–
Tian Hua Fen	16.41	–
Sang Ji Shend	16.73	–
Tu Fu Ling	16.86	–
Ma Huang	16.92	–
Wu Jia Pi	17.26	–
Ji Xue Teng	18.31	–
Jin Yin Hua	18.84	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	19.11	–
He Shou Wu	19.26	–
Jin Qian Cao	19.99	–
Bai Jiang Cao	20.21	–
Gan Cao	20.28	–
Huang Lian	20.71	–
(Fen) Bi Xie	20.90	–
Bai Shao Yao	21.09	–
Lian Qiao	21.15	–
Suan Zao Ren	21.50	–
She Gan	21.52	–
Ban Lan Gen	21.64	–
Gua Lou	21.68	–
He Huan Pi	21.89	–
Xuan Fu Hua	22.39	–
Ma Huang Gen	22.94	–
Gou Teng	23.00	–
Wu Wei Zi	23.10	–
Zhi Ke	23.58	–
Mang Xiao	23.81	–
Di Gu Pi	24.02	–
Mu Zei	24.80	–
Fu Ling	25.10	–
Lu Gen	25.14	–
Rou Gui	25.35	–
Zhi Gan Cao	25.59	–
Bu Gu Zhi	25.60	–
Ren Shen	25.99	–
Yuan Zhi	26.03	–
Sheng Jiang	26.28	–
Huo Ma Ren	26.64	–
Cang Zhu	26.77	–
Zhe Bei Mu	27.78	–
Gou Qi Zi	28.05	–
Ding Xiang	28.21	–
Gui Zhi	28.46	–
Yu Xing Cao	28.94	–
Nü Zhen Zi	29.60	–
Pu Gong Ying	30.05	–
Yi Yi Ren	30.29	–
Ge Gen	30.92	–
Sang Bai Pi	30.98	–
Chuang Mu Xiang	31.52	–
Tao Ren	32.06	–
Sang Ye	32.14	–
Hong Hua	32.38	–
Chen Pi	34.18	–
Wu Zhu Yu	35.24	–
Bai Zi Ren	35.27	–
Chi Shao (Yao)	35.36	–
Ze Lan	35.40	–
Ze Xie	35.73	–
Tai Zi Shen	35.89	–
Sang Zhi	36.45	–
Xuan Shen	37.41	–
Ban Xia (Jiang)	37.69	–
Ci Wu Jia	37.95	–
Huang Qin	38.34	–
Jie Geng	38.62	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xi Xian Cao	38.70	–
Zhu Ru	38.99	–
Cang Er Zi	39.08	–
Ban Zhi Lian	39.18	–
Xie Bai	39.60	–
Yu Zhu	39.68	–
Fo Shou	39.97	–
Dang Gui	41.33	–
Zhi Shi	41.43	–
Xian Mao	41.46	–
Fu Xiao Mai	41.50	–
Yi Mu Cao	41.52	–
Shan Yu Rou	42.05	–
Fu Shen	42.22	–
Dang Gui Wei	43.07	–
San Qi	43.36	–
Long Yan Rou	43.72	–
E Zhu	43.83	–
Jiang Huang	45.52	–
Ku Shen	47.23	–
Lai Fu Zi	48.00	–
Hua Shi	49.04	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jing Jie* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62497	62497	0.00	6.97
62498	62498	0.00	7.51
62751	62751	0.00	14.25
62752	62752	0.00	13.89
62785	62785	0.00	5.03
62786	62786	0.00	4.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jiu Da Huang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	61389-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jiu Da Huang; Rhei radix et rhizoma praeparata

Special notes

When selecting the *Jiu Da Huang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jiu Da Huang	1	0	0

Second-stage model

For differentiation of the substance/substance group *Jiu Da Huang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jiu Da Huang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jiu Da Huang	G561HS013SW1	62821	40	from supplier
PhytoComm	Jiu Da Huang	G561HS013SW1	62822	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Jiu Da Huang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Jiu Da Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jiu Da Huang	G561HS013SW1	62821†	20
PhytoComm	Jiu Da Huang	G561HS013SW1	62822†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Jiu Da Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jiu Da Huang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jiu Da Huang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	0	857

The substance/substance group *Jiu Da Huang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jiu Da Huang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	15.91	–
Xuan Shen	30.46	–
Guang Huo Xiang	33.62	–
Hua Shi	47.38	–
Ding Xiang	48.69	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jiu Da Huang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62821	62821	0.00	15.95
62822	62822	0.00	15.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ju Hua**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002256-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ju Hua; Chrysanthemi flos

Special notes

When selecting the *Ju Hua* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ju Hua	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ju Hua* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ju Hua*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ju Hua	G064H1231822	62775	40	from supplier
PhytoComm	Ju Hua	G064H1231822	62776	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ju Hua*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ju Hua*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ju Hua	G064H1231822	62775†	20
PhytoComm	Ju Hua	G064H1231822	62776†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Ju Hua*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ju Hua	G064H1231421	2
PhytoComm	Ju Hua	G064H1231421	1
PhytoComm	Ju Hua	G064H1231521	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ju Hua* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ju Hua* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	2	0	4	851

The substance/substance group *Ju Hua* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8560 % (> 99.2695 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ju Hua* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chen Pi	16.70	–
Fang Feng	17.67	–
Xiang Fu	17.73	–
Chuan Niu Xi	18.60	–
(Shi) Chang Pu	20.05	–
Shan Yu Rou	23.50	–
Dang Gui	25.11	–
Gua Lou	25.44	–
Mang Xiao	25.90	–
Wu Yao	26.24	–
Bing Lang	26.49	–
Gu Sui Bu	26.69	–
He Huan Pi	28.49	–
Sha Ren	28.76	–
Huang Qi	28.98	–
Chuan Lian Zi	29.36	–
Zhi Gan Cao	30.46	–
Ba Ji Tian	31.39	–
He Shou Wu	32.30	–
Wu Zhu Yu	32.66	–
Chi Shao (Yao)	32.77	–
Xu Duan	34.06	–
Hong Jing Tian	34.25	–
Ku Shen	34.45	–
Yi Mu Cao	34.67	–
Bai Zhi	35.80	–
Shan Yao	38.17	–
Sang Ji Shend	38.53	–
Ji Li	38.67	–
Jiang Huang	39.07	–
Tian Hua Fen	39.59	–
Mu Dan Pi	40.04	–
Cang Er Zi	40.33	–
Sang Zhi	40.74	–
Ge Gen	41.08	–
Xiao Hui Xiang	41.15	–
(Huai) Niu Xi	42.09	–
Sang Bai Pi	42.36	–
Di Gu Pi	42.84	–
Du Huo	42.95	–
E Zhu	43.04	–
Zhi Ke	43.55	–
Bai Zhu	43.60	–
Zi Su Zi	43.71	–
Lian Qiao	43.85	–
Xin Yi	43.92	–
Bai Hua She She Cao	44.05	–
Yuan Zhi	44.08	–
Gan Cao	44.12	–
Wu Wei Zi	44.30	–
Pu Gong Ying	44.51	–
Niu Bang Zi	44.69	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jie Geng	45.14	–
(Bai) Dou Kou	45.99	–
Yan Hu Suo	47.44	–
Sang Ye	50.17	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ju Hua* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62775	62775	0.00	16.76
62776	62776	0.00	16.70

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Jue Ming Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002262-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Jue Ming Zi; Cassiae torae semen

Special notes

When selecting the *Jue Ming Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Jue Ming Zi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Jue Ming Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Jue Ming Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Jue Ming Zi	G059H1097021	62927	40	from supplier
PhytoComm	Jue Ming Zi	G059H1097021	62928	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Jue Ming Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Jue Ming Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Jue Ming Zi	G059H1097021	62927†	20
PhytoComm	Jue Ming Zi	G059H1097021	62928†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 1 batches from the substance/substance group *Jue Ming Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Jue Ming Zi	G059H1097221	1
PhytoComm	Jue Ming Zi	G059H1097221	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Jue Ming Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Jue Ming Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	2	0	4	851

The substance/substance group *Jue Ming Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8450 % (> 99.2584 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Jue Ming Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	15.79	–
Guang Huo Xiang	31.13	–
Shu Di (Huang)	35.48	–
Hua Shi	41.15	–
(Sheng) Di Huang	41.25	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Jue Ming Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62927	62927	0.00	15.89
62928	62928	0.00	15.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ku Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60112-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ku Shen; Sophorae flavescentis radix

Special notes

When selecting the *Ku Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ku Shen	2	0	2

Second-stage model

For differentiation of the substance/substance group *Ku Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ku Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ku Shen	G233H0908822	62499	40	from supplier
PhytoComm	Ku Shen	G233H0908822	62500	40	from supplier
PhytoComm	Ku Shen	G233HS180TK1	62795	40	from supplier
PhytoComm	Ku Shen	G233HS180TK1	62796	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ku Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ku Shen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ku Shen	G233H0908822	62499†	20
PhytoComm	Ku Shen	G233H0908822	62500†	20
PhytoComm	Ku Shen	G233HS180TK1	62795†	20
PhytoComm	Ku Shen	G233HS180TK1	62796†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Ku Shen*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ku Shen	G233H0908121	1
Phytocomm	Ku Shen	h0908121	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ku Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ku Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	2	0	2	853

The substance/substance group *Ku Shen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.7287 % (> 99.1438 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ku Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
San Qi	6.68	–
Lian Qiao	7.97	–
Zhe Bei Mu	8.99	–
Jin Yin Hua	10.01	–
Bai Zi Ren	10.61	–
Shan Yao	10.96	–
Ji Li	11.06	–
Jiao Gu Lan	11.31	–
Cang Zhu	11.78	–
Zhi Mu	12.22	–
Pi Pa Ye	12.52	–
Suan Zao Ren	12.68	–
Ling Zhi	13.64	–
Yuan Zhi	14.85	–
Lai Fu Zi	15.05	–
Tian Hua Fen	15.26	–
Gua Lou	16.00	–
Shan Yu Rou	16.34	–
Ren Shen	16.37	–
Xie Bai	16.41	–
Niu Bang Zi	16.78	–
Huo Ma Ren	17.22	–
Hou Po	17.32	–
Wu Wei Zi	17.41	–
Lian Zi	18.29	–
He Huan Pi	18.62	–
Gou Qi Zi	18.69	–
Dang Gui	18.96	–
Bai Shao Yao	19.00	–
Jie Geng	19.04	–
Chuan Xiong	19.11	–
Zhu Ru	19.14	–
Fu Zi	19.29	–
Ren Dong Teng	19.48	–
Mu Zei	19.62	–
Chuang Mu Xiang	20.49	–
Ze Xie	20.96	–
Fu Ling	21.03	–
Chen Pi	21.05	–
Sheng Jiang	21.11	–
Mao Dong Qing	21.26	–
Cang Er Zi	21.32	–
Tai Zi Shen	21.38	–
Yi Mu Cao	21.39	–
Yu Zhu	21.57	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	21.63	–
Mu Gua	22.66	–
Shen Qu	22.87	–
Di Gu Pi	23.08	–
Dan Dou Chi	23.47	–
Ji Xue Teng	23.89	–
(Huai) Niu Xi	23.97	–
Ban Lan Gen	24.27	–
Guang Huo Xiang	24.43	–
Yan Hu Suo	24.49	–
Dang Gui Wei	24.82	–
Xiang Fu	25.12	–
Xiao Hui Xiang	25.22	–
Dan Shen	25.29	–
Tu Fu Ling	26.26	–
Chai Hu	26.84	–
Gu Sui Bu	27.35	–
Gui Zhi	27.62	–
Lu Gen	27.94	–
Ye Jiao Teng	28.12	–
Fu Pen Zi	28.33	–
Yu Jin	28.81	–
Che Qian Zi	29.15	–
Yi Yi Ren	29.61	–
Qing Pi	29.68	–
Gan Cao	29.85	–
Gou Teng	30.05	–
Zhi Gan Cao	30.73	–
Dong Gua Zi	31.00	–
Ci Wu Jia	31.09	–
Huang Qin	31.10	–
Mang Xiao	31.61	–
Tao Ren	31.62	–
Zhu Ling	32.24	–
Fu Xiao Mai	32.36	–
Bai Xian Pi	32.64	–
Rou Gui	33.05	–
Qiang Huo	33.06	–
Ban Xia (Jiang)	33.13	–
Ce Bai Ye	33.24	–
Ma Huang	33.55	–
Yin Yang Huo	33.65	–
Zi Su Zi	33.82	–
Zi Hua Di Ding	34.10	–
Hong Jing Tian	35.11	–
She Gan	35.31	–
Xu Duan	35.56	–
Gan Jiang	35.69	–
Chi Shao (Yao)	36.19	–
Long Yan Rou	36.63	–
Huang Qi	37.13	–
Jiang Huang	37.13	–
Du Zhong	37.15	–
Wu Mei	37.69	–
Sang Zhi	38.88	–
Huang Lian	39.36	–
(Bai) Dou Kou	40.80	–
Bo He	41.22	–
Hong Hua	41.58	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pu Gong Ying	41.70	–
Fo Shou	42.00	–
(Fen) Bi Xie	42.05	–
Sang Ye	42.51	–
(Shi) Chang Pu	42.92	–
Sha Ren	43.94	–
Ze Lan	44.68	–
Ma Huang Gen	44.88	–
Mai Men Dong	46.30	–
Chuan Lian Zi	46.45	–
Huang Bai	47.41	–
Sha Shen (Bei)	48.53	–
Chuan Niu Xi	48.65	–
Qing Hao	49.16	–
E Zhu	50.08	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ku Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62499	62499	0.00	16.82
62500	62500	0.00	16.58
62795	62795	0.00	6.68
62796	62796	0.00	7.40

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Lai Fu Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60067-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Lai Fu Zi; Raphani sativi semen

Special notes

When selecting the *Lai Fu Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Lai Fu Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Lai Fu Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Lai Fu Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Lai Fu Zi	G209H1220821	62491	40	from supplier
PhytoComm	Lai Fu Zi	G209H1220821	62492	40	from supplier
PhytoComm	Lai Fu Zi	G209HS278TN1	63031	40	from supplier
PhytoComm	Lai Fu Zi	G209HS278TN1	63032	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Lai Fu Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Lai Fu Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Lai Fu Zi	G209H1220821	62491†	20
PhytoComm	Lai Fu Zi	G209H1220821	62492†	20
PhytoComm	Lai Fu Zi	G209HS278TN1	63031†	20
PhytoComm	Lai Fu Zi	G209HS278TN1	63032†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 1 *Apo-Ident* customers from 3 batches from the substance/substance group *Lai Fu Zi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Herbasin	Lai Fu Zi	g209h1220321	1
Phytocomm	Lai Fu Zi	G209H1220221	1
Phytocomm	Lai Fu Zi	G209H1220321	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Lai Fu Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Lai Fu Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	159	1	24 439
Type B	3	73	7	12 394
Type C	0	0	3	854

The substance/substance group *Lai Fu Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9982 % (> 99.9832 %)	99.3750 % (> 97.5000 %)
Type B	99.9667 % (> 99.9368 %)	91.2500 % (> 87.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Lai Fu Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Shao Yao	4.18	–
He Huan Pi	4.40	–
Di Gu Pi	6.16	–
Bai Zi Ren	6.22	–
Shen Qu	6.46	–
Sheng Jiang	6.80	–
Tai Zi Shen	7.09	–
Ci Wu Jia	7.16	–
Zhu Ru	7.93	–
Tao Ren	7.94	–
Ji Li	8.49	–
Zhe Bei Mu	8.65	–
Gui Zhi	8.72	–
Lian Zi	9.20	–
Huo Ma Ren	9.35	–
Mu Zei	9.65	–
Gua Lou	10.83	–
Tian Hua Fen	10.94	–
Ye Jiao Teng	11.12	–
Fu Xiao Mai	11.43	–
Ling Zhi	11.60	–
Ze Xie	12.14	–
Rou Gui	13.64	–
Gou Teng	13.72	–
Yi Yi Ren	13.78	–
Pi Pa Ye	13.95	–
Ban Xia (Jiang)	14.09	–
Shan Yao	14.40	–
Yuan Zhi	14.88	–
Gu Sui Bu	14.92	–
Fo Shou	15.28	–
Fu Ling	15.43	–
Ren Dong Teng	15.65	–
Bai Xian Pi	15.89	–
Lu Gen	16.99	–
Lian Qiao	17.13	–
Chuan Xiong	17.33	–
Suan Zao Ren	17.43	–
Ji Xue Teng	17.77	–
Jin Yin Hua	18.49	–
Tu Fu Ling	19.29	–
Sang Zhi	20.01	–
Chen Pi	20.04	–
Yu Jin	20.41	–
Cang Zhu	21.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Zi	21.90	–
Ma Huang Gen	22.01	–
Jie Geng	22.91	–
Dang Gui Wei	23.04	–
Zhi Mu	23.32	–
Dang Gui	23.51	–
Fu Pen Zi	23.53	–
San Qi	23.59	–
Chuan Lian Zi	23.74	–
Ce Bai Ye	24.21	–
Yu Zhu	24.45	–
Zhu Ling	24.62	–
Yan Hu Suo	24.92	–
(Fen) Bi Xie	25.11	–
Jiao Gu Lan	25.44	–
Bai He	25.47	–
Ban Lan Gen	25.82	–
Dan Shen	25.95	–
Bai Zhu	26.06	–
Dan Dou Chi	27.07	–
She Gan	27.26	–
Zhi Ke	27.55	–
Sha Shen (Bei)	27.78	–
Ren Shen	27.78	–
Shan Yu Rou	28.01	–
Hou Po	28.36	–
Mao Dong Qing	28.42	–
Mang Xiao	28.72	–
Fu Shen	29.47	–
Che Qian Zi	29.86	–
Chai Hu	29.96	–
Guang Huo Xiang	30.69	–
Yin Yang Huo	30.80	–
Mu Gua	32.10	–
Ku Shen	32.86	–
Cang Er Zi	33.11	–
Gou Qi Zi	33.15	–
Ma Huang	33.27	–
Huang Qin	33.36	–
E Zhu	33.67	–
Wu Wei Zi	34.01	–
Hong Jing Tian	35.81	–
Mai Ya	35.86	–
Gan Cao	36.70	–
Qiang Huo	37.42	–
Gan Jiang	39.27	–
Chuang Mu Xiang	39.30	–
Long Yan Rou	39.35	–
Zi Su Zi	39.66	–
Xie Bai	41.69	–
Jiang Huang	42.65	–
Huang Qi	43.25	–
Huang Lian	44.58	–
Zi Hua Di Ding	45.19	–
Jing Jie	46.13	–
Qing Pi	46.44	–
Da Zao	46.53	–
Chuan Mu Tong	47.85	–
Long Dan (Cao)	48.07	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Zhi Lian	49.35	–
Qin Jiao	49.40	–
Sha Ren	49.65	–
Zhi Gan Cao	49.93	–
Du Zhong	49.97	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Lai Fu Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62491	62491	0.00	20.02
62492	62492	0.00	20.01
63031	63031	0.00	4.40
63032	63032	0.00	4.18

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Lian Qiao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60146-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Lian Qiao; Forsythiae fructus

Special notes

When selecting the *Lian Qiao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Lian Qiao	3	0	3

Second-stage model

For differentiation of the substance/substance group *Lian Qiao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Lian Qiao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Lian Qiao	G105H1132821	62337	40	from supplier
PhytoComm	Lian Qiao	G105H1132821	62338	40	from supplier
PhytoComm	Lian Qiao	G105HS252RW1	62685	40	from supplier
PhytoComm	Lian Qiao	G105HS252RW1	62686	40	from supplier
PhytoComm	Lian Qiao	G105H1132021	63005	40	from supplier
PhytoComm	Lian Qiao	G105H1132021	63006	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Lian Qiao*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Lian Qiao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Lian Qiao	G105H1132821	62337†	20
PhytoComm	Lian Qiao	G105H1132821	62338†	20
PhytoComm	Lian Qiao	G105HS252RW1	62685†	20
PhytoComm	Lian Qiao	G105HS252RW1	62686†	20
PhytoComm	Lian Qiao	G105H1132021	63005†	20
PhytoComm	Lian Qiao	G105H1132021	63006†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Lian Qiao*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Lian Qiao	G105H1132221	1
PhytoComm	Lian Qiao	G105H1132321	1
Phytocomm	Lian Qiao	G105H1132521	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Lian Qiao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Lian Qiao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	3	120	0	12 354
Type C	0	0	3	854

The substance/substance group *Lian Qiao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	99.9857 % (> 99.9558 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Lian Qiao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Li	4.93	–
Ling Zhi	6.13	–
Shan Yao	6.59	–
Zhe Bei Mu	8.25	–
Bai Zi Ren	9.48	–
Lian Zi	10.07	–
Pi Pa Ye	10.76	–
Suan Zao Ren	10.98	–
Niu Bang Zi	11.19	–
Jiao Gu Lan	11.50	–
Chuan Xiong	11.70	–
Shen Qu	12.76	–
Fu Zi	13.22	–
Cang Zhu	13.29	–
Mao Dong Qing	13.45	–
Hou Po	13.54	–
Tian Hua Fen	13.67	–
He Huan Pi	14.22	–
Ze Xie	14.26	–
Fu Ling	15.13	–
Jin Yin Hua	15.18	–
Guang Huo Xiang	15.61	–
Huo Ma Ren	16.04	–
San Qi	16.09	–
Ku Shen	16.22	–
Yuan Zhi	16.22	–
Ren Dong Teng	16.33	–
Gua Lou	16.54	–
Chai Hu	16.60	–
Dan Shen	16.64	–
Bai Shao Yao	16.69	–
Mu Zei	17.09	–
Gu Sui Bu	17.54	–
Chen Pi	17.65	–
Dan Dou Chi	17.81	–
Sheng Jiang	17.91	–
Tai Zi Shen	18.12	–
Yu Jin	18.55	–
Tu Fu Ling	19.60	–
Xie Bai	19.68	–
Wu Wei Zi	20.14	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Pen Zi	20.40	–
(Huai) Niu Xi	20.68	–
Yan Hu Suo	20.92	–
She Gan	21.05	–
Bai Xian Pi	21.29	–
Yu Zhu	21.36	–
Zi Su Zi	21.55	–
Di Gu Pi	21.71	–
Zhu Ru	21.90	–
Che Qian Zi	22.14	–
Ban Lan Gen	22.20	–
Gou Qi Zi	22.91	–
Ce Bai Ye	23.01	–
Dang Gui	23.24	–
Gan Cao	23.25	–
Gui Zhi	23.36	–
Zhu Ling	23.53	–
Zhi Gan Cao	23.73	–
Lu Gen	23.83	–
Ren Shen	24.95	–
Lai Fu Zi	25.08	–
Xiang Fu	25.10	–
Ji Xue Teng	25.15	–
Gou Teng	25.16	–
Chuang Mu Xiang	25.66	–
Zhi Ke	25.68	–
Ye Jiao Teng	25.70	–
Qiang Huo	25.80	–
Shan Yu Rou	26.13	–
Qing Pi	26.26	–
Xiao Hui Xiang	26.26	–
Dang Gui Wei	26.31	–
Yin Yang Huo	26.39	–
Chi Shao (Yao)	26.87	–
Jie Geng	27.03	–
Hong Jing Tian	27.28	–
Tao Ren	28.41	–
Yi Mu Cao	28.49	–
Du Zhong	29.40	–
Yi Yi Ren	29.68	–
Mang Xiao	29.70	–
Rou Gui	29.81	–
Ci Wu Jia	30.46	–
Fu Xiao Mai	30.72	–
Zi Hua Di Ding	30.89	–
Zhi Mu	31.26	–
Cang Er Zi	31.63	–
Ban Xia (Jiang)	32.87	–
(Fen) Bi Xie	33.11	–
Ma Huang	33.31	–
Bo He	33.45	–
Ma Huang Gen	34.20	–
Mu Gua	34.89	–
Huang Qin	35.32	–
Huang Lian	35.65	–
Chuan Lian Zi	36.31	–
Bai Zhi	37.03	–
Sang Zhi	37.33	–
Huang Bai	37.36	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Du Huo	37.38	–
Huang Qi	37.53	–
Jing Jie	37.83	–
Sha Ren	37.96	–
(Shi) Chang Pu	38.13	–
Ju Hua	40.13	–
Qing Hao	40.24	–
Hong Hua	40.44	–
Long Yan Rou	40.54	–
Fo Shou	41.46	–
Gan Jiang	41.75	–
Pu Gong Ying	41.78	–
Wu Mei	43.03	–
Jiang Huang	43.23	–
Dong Gua Zi	43.32	–
Qin Jiao	43.46	–
Wu Yao	43.69	–
Fang Feng	44.23	–
(Bai) Dou Kou	44.97	–
Sang Ye	45.80	–
Xu Duan	46.30	–
Chuan Niu Xi	46.42	–
Ze Lan	47.32	–
Bai Jiang Cao	47.83	–
E Zhu	49.31	–
Fu Shen	49.43	–
Chuan Mu Tong	49.71	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Lian Qiao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62337	62337	0.00	18.28
62338	62338	0.00	18.13
62685	62685	0.00	5.51
62686	62686	0.00	4.93
63005	63005	0.00	11.36
63006	63006	0.00	11.19

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances,

thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Lian Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60248-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Lian Zi; Nelumbinis semen

Special notes

When selecting the *Lian Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Lian Zi	1	0	0

Second-stage model

For differentiation of the substance/substance group *Lian Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Lian Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Lian Zi	G171HS312SK1	62567	40	from supplier
PhytoComm	Lian Zi	G171HS312SK1	62568	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Lian Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Lian Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Lian Zi	G171HS312SK1	62567†	20
PhytoComm	Lian Zi	G171HS312SK1	62568†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Lian Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Lian Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Lian Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	2	40	0	12 435
Type C	0	0	0	857

The substance/substance group *Lian Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9905 % (> 99.9608 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Lian Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Suan Zao Ren	6.36	–
Ling Zhi	7.20	–
Ji Li	7.62	–
Shen Qu	9.59	–
Huo Ma Ren	10.03	–
Ren Dong Teng	11.17	–
Tu Fu Ling	11.33	–
He Huan Pi	11.35	–
Fu Ling	11.36	–
Bai Shao Yao	11.44	–
She Gan	11.94	–
Bai Xian Pi	12.47	–
Bai Zi Ren	12.83	–
Sheng Jiang	12.89	–
Zhe Bei Mu	13.39	–
Tai Zi Shen	13.71	–
Tao Ren	13.92	–
Lu Gen	14.82	–
Shan Yao	14.92	–
Ce Bai Ye	14.95	–
Gu Sui Bu	14.95	–
Lian Qiao	15.10	–
Gua Lou	15.24	–
Gui Zhi	15.25	–
Ze Xie	15.28	–
Mu Zei	15.50	–
Dan Dou Chi	15.61	–
Pi Pa Ye	15.62	–
Chai Hu	15.66	–
Ye Jiao Teng	15.99	–
Chuan Xiong	16.09	–
Tian Hua Fen	16.47	–
Yu Jin	17.06	–
Guang Huo Xiang	17.14	–
Ci Wu Jia	17.16	–
Chen Pi	17.23	–
Mao Dong Qing	17.30	–
Gou Teng	17.69	–
Fu Zi	17.75	–
Di Gu Pi	17.99	–
Rou Gui	18.48	–
Yuan Zhi	18.90	–
Zhu Ru	18.95	–
(Fen) Bi Xie	19.14	–
Jiao Gu Lan	20.02	–
Zhu Ling	20.30	–
Yi Yi Ren	20.36	–
Hou Po	20.67	–
Dan Shen	20.90	–
Jin Yin Hua	21.05	–
Fu Pen Zi	21.09	–
Fu Xiao Mai	21.50	–
Ban Lan Gen	21.77	–
Ji Xue Teng	21.87	–
San Qi	22.02	–
Yu Zhu	22.09	–
Hong Jing Tian	22.17	–
Lai Fu Zi	22.52	–
Cang Zhu	22.57	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Che Qian Zi	22.90	–
Dang Gui Wei	23.75	–
Ma Huang Gen	23.93	–
Fo Shou	23.94	–
Yin Yang Huo	24.75	–
Gan Cao	25.22	–
Ban Xia (Jiang)	25.55	–
Qiang Huo	27.43	–
Chuang Mu Xiang	27.65	–
Yan Hu Suo	27.76	–
Zhi Ke	28.04	–
Jie Geng	29.05	–
Jing Jie	29.36	–
Ma Huang	29.93	–
Dang Gui	29.93	–
Wu Wei Zi	30.49	–
Zi Hua Di Ding	31.31	–
Xie Bai	31.57	–
Gou Qi Zi	31.90	–
Ren Shen	31.92	–
Shan Yu Rou	32.89	–
Zhi Gan Cao	33.18	–
Mang Xiao	33.57	–
Bo He	35.64	–
Fu Shen	36.38	–
Qing Pi	36.77	–
Sha Ren	37.63	–
Ku Shen	38.30	–
Du Zhong	38.72	–
Zhi Mu	39.04	–
Mu Gua	39.23	–
Sang Zhi	39.34	–
Huang Lian	39.54	–
Huang Qin	39.62	–
Huang Bai	39.66	–
Long Yan Rou	40.66	–
Chi Shao (Yao)	45.43	–
Bai Zhu	49.05	–
Cang Er Zi	49.63	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Lian Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62567	62567	0.00	6.36
62568	62568	0.00	7.20

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ling Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60187-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ling Zhi; Ganoderma

Special notes

When selecting the *Ling Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ling Zhi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Ling Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ling Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ling Zhi	G109HS355RP1	62453	40	from supplier
PhytoComm	Ling Zhi	G109HS355RP1	62454	40	from supplier
PhytoComm	Ling Zhi	G109HS355TG1	62941	40	from supplier
PhytoComm	Ling Zhi	G109HS355TG1	62942	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ling Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ling Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ling Zhi	G109HS355RP1	62453†	20
PhytoComm	Ling Zhi	G109HS355RP1	62454†	20
PhytoComm	Ling Zhi	G109HS355TG1	62941†	20
PhytoComm	Ling Zhi	G109HS355TG1	62942†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Ling Zhi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Ling Zhi	G109F-B2405221	1
Phytocomm	Ling Zhi	g109fb2405321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ling Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ling Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	159	1	24 439
Type B	5	79	1	12 392
Type C	0	0	2	855

The substance/substance group *Ling Zhi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	99.3750 % (> 97.5000 %)
Type B	99.9667 % (> 99.9368 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ling Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Li	4.86	–
Ren Dong Teng	5.31	–
Lian Zi	5.79	–
Ce Bai Ye	6.00	–
Bai Xian Pi	7.09	–
Yu Jin	7.13	–
Dan Dou Chi	7.22	–
Chai Hu	7.71	–
He Huan Pi	8.27	–
Shen Qu	8.43	–
Hou Po	8.50	–
Bai Zi Ren	9.79	–
Lian Qiao	9.85	–
Zhe Bei Mu	10.18	–
Mao Dong Qing	10.19	–
Huo Ma Ren	10.35	–
Bai Shao Yao	10.88	–
Che Qian Zi	11.48	–
Zhu Ling	11.62	–
Guang Huo Xiang	11.63	–
Sheng Jiang	11.66	–
Tu Fu Ling	11.88	–
Tian Hua Fen	11.90	–
Ze Xie	12.07	–
Gu Sui Bu	12.41	–
Fu Ling	12.79	–
Fu Zi	12.98	–
She Gan	13.15	–
Tai Zi Shen	13.22	–
Shan Yao	13.63	–
Yin Yang Huo	13.99	–
Mu Zei	14.00	–
Ye Jiao Teng	14.09	–
Suan Zao Ren	14.25	–
Fu Pen Zi	15.01	–
Pi Pa Ye	15.34	–
Gui Zhi	15.47	–
Zhu Ru	15.71	–
Di Gu Pi	15.79	–
Yan Hu Suo	16.20	–
Dan Shen	16.42	–
Chuan Xiong	16.50	–
Jiao Gu Lan	16.85	–
Gua Lou	17.05	–
(Fen) Bi Xie	17.24	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Teng	17.30	–
Qiang Huo	17.59	–
Lai Fu Zi	17.73	–
Lu Gen	18.07	–
Tao Ren	18.15	–
Fu Xiao Mai	18.66	–
Yi Yi Ren	19.58	–
Yuan Zhi	19.63	–
Ji Xue Teng	19.70	–
Rou Gui	19.71	–
Jing Jie	19.76	–
Gan Cao	20.05	–
Jin Yin Hua	21.17	–
Chen Pi	21.41	–
Hong Jing Tian	21.61	–
Zi Hua Di Ding	21.68	–
San Qi	21.79	–
Ci Wu Jia	21.88	–
Cang Zhu	22.31	–
Du Zhong	22.42	–
Ma Huang Gen	22.59	–
Qing Pi	23.39	–
Ban Lan Gen	24.56	–
Ban Xia (Jiang)	24.80	–
Wu Wei Zi	25.45	–
Zhi Ke	26.10	–
Yu Zhu	26.52	–
Ma Huang	26.81	–
Sha Ren	26.94	–
Zhi Gan Cao	27.12	–
Bo He	27.38	–
Dang Gui Wei	27.85	–
Huang Bai	28.04	–
Ren Shen	28.75	–
Dang Gui	29.28	–
Chuang Mu Xiang	29.37	–
Ku Shen	29.64	–
Jie Geng	30.33	–
Qing Hao	30.54	–
Fo Shou	31.35	–
Huang Lian	31.40	–
Xie Bai	32.42	–
Shan Yu Rou	32.74	–
Mang Xiao	33.62	–
Gou Qi Zi	33.87	–
Zhi Mu	37.49	–
Huang Qin	38.60	–
Fu Shen	39.20	–
Bai Jiang Cao	40.01	–
Mu Gua	40.58	–
Sang Zhi	41.44	–
Chi Shao (Yao)	44.30	–
Cang Er Zi	44.40	–
Long Yan Rou	47.30	–
Hong Hua	50.01	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ling Zhi* is separated from critical neighbours in a second-stage model, all

the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62453	62453	0.00	7.13
62454	62454	0.00	5.31
62941	62941	0.00	5.18
62942	62942	0.00	4.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Long Dan (Cao)**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60227-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Long Dan (Cao); Gentianae scabrae radix

Special notes

When selecting the *Long Dan (Cao)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Long Dan (Cao)	2	0	2

Second-stage model

For differentiation of the substance/substance group *Long Dan (Cao)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Long Dan (Cao)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Long Dan (Cao)	G112H1705821	62531	40	from supplier
PhytoComm	Long Dan (Cao)	G112H1705821	62532	40	from supplier
PhytoComm	Long Dan (Cao)	G112H1705921	62871	40	from supplier
PhytoComm	Long Dan (Cao)	G112H1705921	62872	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Long Dan (Cao)*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Long Dan (Cao)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Long Dan (Cao)	G112H1705821	62531†	20
PhytoComm	Long Dan (Cao)	G112H1705821	62532†	20
PhytoComm	Long Dan (Cao)	G112H1705921	62871†	20
PhytoComm	Long Dan (Cao)	G112H1705921	62872†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Long Dan (Cao)*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Long Dan (Cao)	g112h1705022	1
Phytocomm	Long Dan (Cao)	G112H1705321	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Long Dan (Cao)* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Long Dan (Cao)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	1	3	1	852

The substance/substance group *Long Dan (Cao)* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.9612 % (> 99.3747 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Long Dan (Cao)* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Qin Jiao	8.42	–
Huang Qi	14.37	–
Da Zao	16.93	–
Bai He	17.83	–
Mai Ya	17.89	–
Bai Zhu	18.77	–
Jie Geng	19.02	–
Dang Gui	19.85	–
Zhi Gan Cao	20.36	–
Mang Xiao	21.19	–
Chuan Mu Tong	21.39	–
Sang Zhi	22.64	–
Chuan Lian Zi	23.07	–
Lai Fu Zi	23.30	–
Sha Shen (Bei)	23.77	–
E Zhu	24.16	–
Tian Hua Fen	25.32	–
Chuan Niu Xi	25.91	–
Ji Li	26.53	–
Di Gu Pi	26.76	–
Mi Huan Jun	26.97	–
Mu Gua	27.22	–
(Huai) Niu Xi	28.01	–
Chi Shao (Yao)	29.99	–
Zi Su Zi	30.16	–
Bing Lang	31.38	–
Bai Zhi	32.78	–
Yuan Zhi	34.93	–
Jiang Huang	35.15	–
Gan Jiang	36.27	–
Yan Hu Suo	37.16	–
He Huan Pi	37.77	–
Sha Ren	38.77	–
Chen Pi	39.59	–
Gua Lou	40.12	–
Mu Dan Pi	40.93	–
(Shi) Chang Pu	41.34	–
Shan Yao	41.97	–
Lian Qiao	43.00	–
Ban Zhi Lian	44.15	–
Ba Ji Tian	44.89	–
Cang Er Zi	45.09	–
Tu Fu Ling	45.92	–
Lian Zi	48.99	–
(Bai) Dou Kou	49.53	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Shao Yao	49.54	–
Fang Feng	49.79	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Long Dan (Cao)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62531	62531	0.00	17.89
62532	62532	0.00	16.93
62871	62871	0.00	8.42
62872	62872	0.00	9.14

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Long Yan Rou**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60463-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Long Yan Rou; Euphoriae longanae arillus

Special notes

When selecting the *Long Yan Rou* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Long Yan Rou	1	0	1

Second-stage model

For differentiation of the substance/substance group *Long Yan Rou* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Long Yan Rou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Long Yan Rou	G144HS324SQ1	62727	40	from supplier
PhytoComm	Long Yan Rou	G144HS324SQ1	62728	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Long Yan Rou*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Long Yan Rou*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Long Yan Rou	G144HS324SQ1	62727†	20
PhytoComm	Long Yan Rou	G144HS324SQ1	62728†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Long Yan Rou*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Long Yan Rou	G144H1708221	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Long Yan Rou* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Long Yan Rou* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Long Yan Rou* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Long Yan Rou* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chen Pi	11.51	–
Cang Zhu	11.98	–
Yu Zhu	12.11	–
Dang Gui	12.28	–
Gou Qi Zi	12.36	–
Gua Lou	12.87	–
Ren Shen	13.53	–
Jie Geng	14.89	–
Fo Shou	15.03	–
Chuan Xiong	16.57	–
Dang Gui Wei	16.74	–
Bai Shao Yao	17.64	–
Ye Jiao Teng	17.77	–
Jin Yin Hua	18.61	–
Zhe Bei Mu	18.69	–
Di Gu Pi	19.07	–
Ze Xie	19.68	–
Mu Gua	19.90	–
Ma Huang	20.04	–
Yuan Zhi	20.27	–
Zhi Mu	20.85	–
Tao Ren	21.02	–
Mu Zei	21.06	–
Chuang Mu Xiang	21.19	–
San Qi	21.79	–
Hong Jing Tian	21.89	–
Ban Lan Gen	22.23	–
Tai Zi Shen	22.31	–
Tian Hua Fen	22.51	–
Shen Qu	22.89	–
Ren Dong Teng	24.15	–
Ling Zhi	24.28	–
Lu Gen	24.30	–
Ban Xia (Jiang)	24.40	–
Zhi Ke	24.46	–
Gou Teng	24.71	–
He Huan Pi	24.73	–
Tu Fu Ling	24.79	–
Gu Sui Bu	24.87	–
Pi Pa Ye	24.90	–
Fu Pen Zi	25.00	–
Rou Gui	25.01	–
Dan Dou Chi	25.13	–
Lian Zi	25.25	–
Suan Zao Ren	25.63	–
Xie Bai	25.79	–
Shan Yao	26.55	–
Huang Qin	26.92	–
Gui Zhi	27.05	–
Ci Wu Jia	27.26	–
Lian Qiao	27.33	–
Dan Shen	27.49	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yu Rou	27.51	–
Bai Xian Pi	27.88	–
Gan Cao	27.96	–
Bai Zi Ren	27.98	–
Ji Li	28.93	–
Fu Xiao Mai	29.09	–
Ku Shen	29.09	–
Lai Fu Zi	29.12	–
Jiao Gu Lan	29.44	–
Ji Xue Teng	29.47	–
Yin Yang Huo	29.60	–
Mang Xiao	29.83	–
Sheng Jiang	30.91	–
Fu Ling	31.23	–
Huo Ma Ren	31.28	–
Fu Zi	31.34	–
Sang Zhi	31.37	–
She Gan	32.14	–
Hou Po	32.14	–
Hong Hua	32.18	–
Chi Shao (Yao)	33.36	–
Ma Huang Gen	33.62	–
Mai Men Dong	33.72	–
Wu Wei Zi	33.90	–
(Fen) Bi Xie	33.99	–
Chai Hu	34.53	–
Zi Hua Di Ding	35.09	–
Chuan Lian Zi	35.30	–
Zhu Ru	35.71	–
Yan Hu Suo	35.76	–
Mao Dong Qing	36.34	–
Fu Shen	37.10	–
Yu Jin	37.27	–
Huang Qi	37.39	–
Guang Huo Xiang	37.46	–
Zhi Gan Cao	37.96	–
Ce Bai Ye	38.11	–
Jing Jie	39.33	–
Che Qian Zi	39.56	–
Qiang Huo	39.80	–
Yi Yi Ren	40.97	–
Bo He	42.57	–
Zhu Ling	43.67	–
Sha Ren	44.65	–
Huang Lian	46.10	–
Bai Zhu	46.13	–
Huang Bai	47.64	–
Bai He	50.12	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Long Yan Rou* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62727	62727	0.00	11.51
62728	62728	0.00	11.87

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Lu Gen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10004528-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Lu Gen; Phragmitis communis rhizoma

Special notes

When selecting the *Lu Gen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Lu Gen	1	0	2

Second-stage model

For differentiation of the substance/substance group *Lu Gen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Lu Gen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Lu Gen	G189HS348TG1	62839	40	from supplier
PhytoComm	Lu Gen	G189HS348TG1	62840	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Lu Gen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Lu Gen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Lu Gen	G189HS348TG1	62839†	20
PhytoComm	Lu Gen	G189HS348TG1	62840†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Lu Gen*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Lu Gen	G189H2014221	1
Phytocomm	Lu Gen	G189H2014521	1
PhytoComm	Lu Gen	G189H2014521	2

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Lu Gen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Lu Gen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	4	79	1	24 516
Type B	2	37	3	12 435
Type C	0	0	4	853

The substance/substance group *Lu Gen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9857 % (> 99.9708 %)	98.7500 % (> 95.0000 %)
Type B	99.9857 % (> 99.9560 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Lu Gen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Xue Teng	3.57	–
He Huan Pi	5.80	–
Ye Jiao Teng	6.53	–
Ren Dong Teng	6.92	–
Mu Zei	6.92	–
(Fen) Bi Xie	7.31	–
Gou Teng	7.36	–
Rou Gui	7.80	–
Gui Zhi	7.80	–
Huo Ma Ren	8.01	–
Bai Shao Yao	9.08	–
Gu Sui Bu	9.50	–
Fo Shou	9.54	–
Ma Huang Gen	9.63	–
She Gan	9.75	–
Bai Xian Pi	10.53	–
Fu Ling	10.69	–
Tai Zi Shen	10.99	–
Tao Ren	11.29	–
Sheng Jiang	13.56	–
Suan Zao Ren	13.58	–
Ling Zhi	13.69	–
Shen Qu	13.71	–
Lian Zi	14.00	–
Ban Xia (Jiang)	14.04	–
Ci Wu Jia	14.16	–
Tu Fu Ling	14.94	–
Yi Yi Ren	15.29	–
Dan Dou Chi	15.37	–
Zhu Ling	16.41	–
Zhe Bei Mu	16.43	–
Dan Shen	16.92	–
Lai Fu Zi	16.97	–
Chuan Xiong	17.84	–
Ji Li	18.56	–
Ce Bai Ye	18.66	–
Yu Jin	18.70	–
Fu Xiao Mai	18.81	–
Ma Huang	19.11	–
Guang Huo Xiang	19.16	–
Gua Lou	19.16	–
Jin Yin Hua	19.20	–
Di Gu Pi	19.25	–
Yin Yang Huo	19.55	–
Yan Hu Suo	19.57	–
Ze Xie	19.69	–
Tian Hua Fen	20.28	–
Bai Zi Ren	20.50	–
Lian Qiao	20.64	–
Fu Shen	21.28	–
Chen Pi	21.46	–
Pi Pa Ye	21.68	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Hong Jing Tian	22.31	–
Fu Pen Zi	22.67	–
Fu Zi	23.25	–
Shan Yao	23.62	–
Zhu Ru	23.66	–
Ban Lan Gen	24.15	–
Yu Zhu	24.42	–
Mao Dong Qing	24.47	–
Chai Hu	25.74	–
Yuan Zhi	25.79	–
Jiao Gu Lan	26.15	–
Gan Cao	26.20	–
Hou Po	26.62	–
Cang Zhu	26.91	–
Zhi Ke	26.95	–
Jing Jie	28.59	–
Dang Gui Wei	28.72	–
Ren Shen	29.02	–
Qiang Huo	29.18	–
Dang Gui	30.00	–
Jie Geng	30.08	–
Mang Xiao	31.54	–
Zi Hua Di Ding	32.36	–
Huang Qin	32.55	–
Che Qian Zi	32.95	–
Long Yan Rou	33.68	–
Sha Ren	34.04	–
Zhi Mu	34.35	–
Huang Lian	35.62	–
San Qi	35.93	–
Wu Wei Zi	35.97	–
Sang Zhi	37.22	–
Gou Qi Zi	37.85	–
Chuang Mu Xiang	38.50	–
Shan Yu Rou	39.64	–
Bo He	40.17	–
Mu Gua	40.42	–
Zhi Gan Cao	41.34	–
Du Zhong	41.38	–
Chi Shao (Yao)	43.83	–
Qing Pi	44.29	–
Huang Bai	44.60	–
Ku Shen	45.82	–
Xie Bai	48.39	–
Bai Zhu	49.73	–
Hong Hua	49.86	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Lu Gen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62839	62839	0.00	3.57
62840	62840	0.00	4.32

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ma Huang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50283-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ma Huang; Ephedrae herba

Special notes

When selecting the *Ma Huang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ma Huang	2	0	2

Second-stage model

For differentiation of the substance/substance group *Ma Huang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ma Huang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ma Huang	G097HS255SL1	62693	40	from supplier
PhytoComm	Ma Huang	G097HS255SL1	62694	40	from supplier
PhytoComm	Ma Huang	G097HS255TK2	62939	40	from supplier
PhytoComm	Ma Huang	G097HS255TK2	62940	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ma Huang*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ma Huang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ma Huang	G097HS255SL1	62693†	20
PhytoComm	Ma Huang	G097HS255SL1	62694†	20
PhytoComm	Ma Huang	G097HS255TK2	62939†	20
PhytoComm	Ma Huang	G097HS255TK2	62940†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Ma Huang*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ma Huang	G097H1101522	2
PhytoComm	Ma Huang	G097H1101623	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ma Huang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ma Huang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	3	854

The substance/substance group *Ma Huang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ma Huang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	7.27	–
Fu Pen Zi	11.43	–
Bai Shao Yao	11.51	–
Zhi Ke	11.78	–
Ye Jiao Teng	12.11	–
Lu Gen	13.95	–
Hou Po	14.97	–
Gan Cao	15.61	–
Guang Huo Xiang	15.79	–
Ji Xue Teng	15.82	–
Tian Hua Fen	15.98	–
Mu Zei	15.98	–
Ze Xie	16.12	–
Hong Jing Tian	16.21	–
Zi Hua Di Ding	16.45	–
Dan Dou Chi	16.74	–
Ren Dong Teng	16.94	–
Yan Hu Suo	17.02	–
Ban Lan Gen	17.14	–
Fu Zi	17.27	–
Yin Yang Huo	17.61	–
Tao Ren	17.76	–
Pi Pa Ye	17.80	–
Huang Bai	17.85	–
He Huan Pi	18.12	–
Yu Zhu	18.28	–
Hong Hua	18.70	–
Cang Zhu	18.76	–
Huang Qin	18.81	–
Zhi Gan Cao	19.02	–
Tai Zi Shen	19.25	–
Dan Shen	19.30	–
Rou Gui	19.38	–
Lian Qiao	19.45	–
Suan Zao Ren	20.05	–
Gou Qi Zi	20.08	–
Gu Sui Bu	20.48	–
Gua Lou	20.60	–
Gui Zhi	21.06	–
Chuan Xiong	21.35	–
Chen Pi	21.39	–
Dang Gui	22.10	–
Bo He	22.49	–
Zhe Bei Mu	22.77	–
Huo Ma Ren	22.93	–
Jiao Gu Lan	22.98	–
(Fen) Bi Xie	23.33	–
Lai Fu Zi	23.34	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Zi	23.41	–
She Gan	23.78	–
Bai Xian Pi	23.90	–
Sheng Jiang	23.96	–
Shan Yao	24.16	–
Shen Qu	24.42	–
Yuan Zhi	24.55	–
Dang Gui Wei	24.56	–
Ci Wu Jia	24.60	–
Huang Lian	24.96	–
Gou Teng	25.02	–
Di Gu Pi	25.37	–
Chuang Mu Xiang	26.39	–
Ren Shen	26.65	–
Qiang Huo	26.69	–
Tu Fu Ling	26.74	–
Yu Jin	27.25	–
Jie Geng	27.27	–
Ling Zhi	27.29	–
Ji Li	27.56	–
Zhi Mu	27.62	–
Ce Bai Ye	27.82	–
Wu Wei Zi	27.99	–
Chai Hu	28.39	–
Mang Xiao	28.42	–
Jing Jie	28.71	–
Che Qian Zi	28.82	–
Ku Shen	29.19	–
Chi Shao (Yao)	29.42	–
Fo Shou	29.43	–
Long Yan Rou	29.76	–
Sha Ren	29.80	–
Bai Zi Ren	29.89	–
Ban Xia (Jiang)	29.92	–
Ma Huang Gen	30.08	–
Fu Ling	30.40	–
San Qi	31.01	–
Shan Yu Rou	31.38	–
Mao Dong Qing	32.42	–
Zhu Ling	32.49	–
Qing Pi	33.62	–
Sang Ye	34.66	–
Fu Shen	34.66	–
Mu Gua	35.60	–
Yi Yi Ren	36.10	–
Du Zhong	36.35	–
Bai Jiang Cao	36.50	–
Fu Xiao Mai	37.81	–
Zhu Ru	38.70	–
Ban Zhi Lian	39.14	–
Sang Zhi	40.50	–
Cang Er Zi	41.40	–
Xie Bai	41.72	–
Pu Gong Ying	42.41	–
Ze Lan	43.36	–
Qing Hao	45.87	–
Huang Qi	48.60	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ma Huang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62693	62693	0.00	7.91
62694	62694	0.00	7.27
62939	62939	0.00	11.78
62940	62940	0.00	12.36

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ma Huang Gen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60063-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ma Huang Gen; Ephedrae radix

Special notes

When selecting the *Ma Huang Gen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ma Huang Gen	1	0	0

Second-stage model

For differentiation of the substance/substance group *Ma Huang Gen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ma Huang Gen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ma Huang Gen	G278HS393RN1	62261	40	from supplier
PhytoComm	Ma Huang Gen	G278HS393RN1	62262	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ma Huang Gen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ma Huang Gen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ma Huang Gen	G278HS393RN1	62261†	20
PhytoComm	Ma Huang Gen	G278HS393RN1	62262†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ma Huang Gen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ma Huang Gen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ma Huang Gen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	78	2	24 519
Type B	6	35	5	12 431
Type C	1	0	0	856

The substance/substance group *Ma Huang Gen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9976 % (> 99.9827 %)	97.5000 % (> 93.7500 %)
Type B	99.9714 % (> 99.9417 %)	87.5000 % (> 80.0000 %)
Type C	99.8843 % (> 99.3015 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ma Huang Gen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Rou Gui	3.14	–
He Huan Pi	3.63	–
Gou Teng	6.19	–
Fu Shen	8.34	–
Gui Zhi	10.73	–
Yi Yi Ren	10.94	–
Ji Xue Teng	11.62	–
Fu Ling	11.66	–
Sheng Jiang	14.12	–
Mu Zei	17.02	–
Ban Xia (Jiang)	17.40	–
Fo Shou	18.29	–
Fu Xiao Mai	19.63	–
Ye Jiao Teng	19.92	–
Lu Gen	20.51	–
Bai Xian Pi	20.93	–
Di Gu Pi	21.30	–
Huo Ma Ren	21.67	–
Bai Shao Yao	23.26	–
Lai Fu Zi	23.31	–
Lian Zi	23.32	–
Ren Dong Teng	23.95	–
Ci Wu Jia	24.18	–
Ling Zhi	24.35	–
Tao Ren	24.63	–
Tai Zi Shen	24.93	–
Gu Sui Bu	25.52	–
Zhu Ling	26.54	–
Bai Zi Ren	27.04	–
Zhu Ru	27.13	–
Zhe Bei Mu	27.92	–
Tu Fu Ling	28.04	–
Suan Zao Ren	28.46	–
Dan Dou Chi	28.73	–
She Gan	28.77	–
Shen Qu	29.00	–
(Fen) Bi Xie	29.88	–
Ji Li	30.20	–
Mang Xiao	31.84	–
Ze Xie	32.13	–
Yu Jin	32.61	–
Dan Shen	32.91	–
Tian Hua Fen	34.07	–
Lian Qiao	34.24	–
Yan Hu Suo	34.27	–
Chuan Xiong	34.48	–
Chen Pi	35.75	–
Gua Lou	37.61	–
Shan Yao	37.65	–
Pi Pa Ye	37.95	–
Yuan Zhi	38.44	–
Jiao Gu Lan	38.82	–
Yin Yang Huo	38.99	–
Zhi Ke	39.19	–
Ren Shen	41.46	–
Jin Yin Hua	41.89	–
Hong Jing Tian	41.95	–
Mao Dong Qing	42.60	–
Dang Gui	42.66	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Zhu	42.82	–
Chai Hu	43.04	–
Ce Bai Ye	43.32	–
Guang Huo Xiang	43.68	–
Zhi Mu	43.99	–
Cang Zhu	44.24	–
Huang Qin	44.43	–
Ma Huang	44.63	–
Jie Geng	46.08	–
Shan Yu Rou	46.26	–
Fu Pen Zi	47.58	–
Fu Zi	47.98	–
Ban Lan Gen	47.98	–
Qiang Huo	50.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ma Huang Gen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62261	62261	0.00	3.14
62262	62262	0.00	6.00

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mai Men Dong**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60024-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mai Men Dong; Ophiopogonis radix

Special notes

When selecting the *Mai Men Dong* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mai Men Dong	1	0	2

Second-stage model

For differentiation of the substance/substance group *Mai Men Dong* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mai Men Dong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mai Men Dong	G176HS236SH1	62611	40	from supplier
PhytoComm	Mai Men Dong	G176HS236SH1	62612	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Mai Men Dong*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Mai Men Dong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mai Men Dong	G176HS236SH1	62611†	20
PhytoComm	Mai Men Dong	G176HS236SH1	62612†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Mai Men Dong*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mai Men Dong	g176h1118221	1
Phytocomm	Mai Men Dong	G176H1118521	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mai Men Dong* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mai Men Dong* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	4	853

The substance/substance group *Mai Men Dong* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mai Men Dong* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Gua Lou	21.66	–
Fo Shou	21.67	–
Jie Geng	22.64	–
Ye Jiao Teng	24.32	–
Yuan Zhi	25.47	–
Rou Gui	27.16	–
Fu Xiao Mai	28.29	–
Bai Shao Yao	28.61	–
Di Gu Pi	29.26	–
Zhi Mu	30.30	–
Ren Shen	30.90	–
Mang Xiao	31.10	–
Cang Zhu	31.69	–
Shan Yu Rou	31.71	–
Ji Li	32.43	–
He Huan Pi	32.84	–
Jin Yin Hua	32.96	–
Chuan Xiong	33.19	–
Long Yan Rou	33.58	–
Pi Pa Ye	33.71	–
Yu Zhu	34.07	–
Mu Zei	34.56	–
Fu Ling	34.73	–
Tai Zi Shen	35.13	–
Shen Qu	35.26	–
Dang Gui	36.11	–
Chuang Mu Xiang	36.64	–
Lian Zi	36.83	–
Chuan Lian Zi	37.16	–
Bai Xian Pi	37.43	–
Sheng Jiang	38.09	–
Tian Hua Fen	38.51	–
San Qi	38.73	–
Ban Xia (Jiang)	38.77	–
Zhu Ru	39.54	–
Chen Pi	40.00	–
Gui Zhi	40.05	–
Huang Qin	40.11	–
Gou Teng	40.16	–
Fu Shen	40.44	–
Dang Gui Wei	40.55	–
Tao Ren	40.64	–
Gou Qi Zi	40.69	–
Mu Gua	41.20	–
Shan Yao	41.65	–
Ban Lan Gen	41.94	–
Ren Dong Teng	42.30	–
Zhi Ke	42.92	–
Ci Wu Jia	43.18	–
Ze Xie	43.74	–
Xie Bai	44.43	–
Ji Xue Teng	44.48	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang	44.69	–
Bai Zi Ren	45.68	–
Jiao Gu Lan	46.06	–
Yi Yi Ren	46.19	–
Lian Qiao	46.49	–
Ma Huang Gen	46.62	–
Tu Fu Ling	46.92	–
Lai Fu Zi	47.09	–
Ling Zhi	47.21	–
Ku Shen	48.54	–
Suan Zao Ren	49.07	–
Fu Pen Zi	49.35	–
Lu Gen	49.49	–
Zhe Bei Mu	49.85	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mai Men Dong* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62611	62611	0.00	21.67
62612	62612	0.00	21.66

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mai Ya**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60210-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mai Ya; Hordei vulgaris fructus germinatus

Special notes

When selecting the *Mai Ya* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mai Ya	1	0	2

Second-stage model

For differentiation of the substance/substance group *Mai Ya* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mai Ya*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mai Ya	G123H1119821	62535	40	from supplier
PhytoComm	Mai Ya	G123H1119821	62536	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Mai Ya*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Mai Ya*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mai Ya	G123H1119821	62535†	20
PhytoComm	Mai Ya	G123H1119821	62536†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Mai Ya*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mai Ya	g123h1119022	1
Phytocomm	Mai Ya	G123H1119521	1
PhytoComm	Mai Ya	G123H1119821	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mai Ya* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mai Ya* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	1	2	854

The substance/substance group *Mai Ya* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mai Ya* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai He	16.18	–
Chuan Mu Tong	16.33	–
Mi Huan Jun	17.14	–
Sang Zhi	18.26	–
Long Dan (Cao)	18.43	–
Mang Xiao	21.63	–
Chuan Lian Zi	22.09	–
Di Gu Pi	22.29	–
Gan Jiang	22.40	–
Dang Gui	23.39	–
Ban Zhi Lian	24.02	–
Da Zao	25.56	–
E Zhu	25.87	–
Sha Shen (Bei)	26.91	–
Ji Li	30.07	–
Yan Hu Suo	30.43	–
Lai Fu Zi	31.44	–
Tian Hua Fen	33.46	–
Jiang Huang	33.79	–
Qin Jiao	35.16	–
(Huai) Niu Xi	37.05	–
Huang Qi	39.10	–
Sha Ren	39.94	–
Shan Yao	40.17	–
Cang Er Zi	40.24	–
Rou Gui	41.25	–
He Huan Pi	41.75	–
Zi Su Zi	42.12	–
Fu Ling	43.13	–
Chen Pi	43.58	–
(Bai) Dou Kou	43.92	–
Yi Yi Ren	44.07	–
Jie Geng	46.33	–
Chuan Niu Xi	47.81	–
Lian Zi	48.27	–
(Shi) Chang Pu	48.70	–
Bing Lang	48.76	–
Niu Bang Zi	49.17	–
Sheng Jiang	49.72	–
Bai Zhu	50.32	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mai Ya* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62535	62535	0.00	16.18
62536	62536	0.00	16.33

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Man Jing Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60025-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Man Jing Zi; Viticis fructus

Special notes

When selecting the *Man Jing Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Man Jing Zi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Man Jing Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Man Jing Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Man Jing Zi	G250H1505822	62501	40	from supplier
PhytoComm	Man Jing Zi	G250H1505822	62502	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Man Jing Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Man Jing Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Man Jing Zi	G250H1505822	62501†	20
PhytoComm	Man Jing Zi	G250H1505822	62502†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Man Jing Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Man Jing Zi	G250H1505122	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Man Jing Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Man Jing Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Man Jing Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Man Jing Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hua Shi	30.80	–
Mang Xiao	57.21	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Man Jing Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62501	62501	0.00	30.84
62502	62502	0.00	30.80

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mang Xiao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50387-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mang Xiao; Mirabilitum

Special notes

When selecting the *Mang Xiao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mang Xiao	2	0	0

Second-stage model

For differentiation of the substance/substance group *Mang Xiao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mang Xiao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mang Xiao	G160H0627821	62469	40	from supplier
PhytoComm	Mang Xiao	G160H0627821	62470	40	from supplier
PhytoComm	Mang Xiao	G160HS137TH1	62873	40	from supplier
PhytoComm	Mang Xiao	G160HS137TH1	62874	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Mang Xiao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Mang Xiao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mang Xiao	G160H0627821	62469†	20
PhytoComm	Mang Xiao	G160H0627821	62470†	20
PhytoComm	Mang Xiao	G160HS137TH1	62873†	20
PhytoComm	Mang Xiao	G160HS137TH1	62874†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Mang Xiao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mang Xiao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mang Xiao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	150	10	24 440
Type B	0	77	3	12 397
Type C	0	0	0	857

The substance/substance group *Mang Xiao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	93.7500 % (> 91.8750 %)
Type B	100.0000 % (> 99.9403 %)	96.2500 % (> 92.5000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mang Xiao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hua Shi	44.13	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mang Xiao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62469	62469	0.00	44.95
62470	62470	0.00	45.19
62873	62873	0.00	44.41
62874	62874	0.00	44.13

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mao Dong Qing**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60028-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mao Dong Qing; Ilicis pubescendis radix

Special notes

When selecting the *Mao Dong Qing* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mao Dong Qing	2	0	0

Second-stage model

For differentiation of the substance/substance group *Mao Dong Qing* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mao Dong Qing*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mao Dong Qing	G319HS063RN1	62417	40	from supplier
PhytoComm	Mao Dong Qing	G319HS063RN1	62418	40	from supplier
PhytoComm	Mao Dong Qing	G319HS063SK1	62619	40	from supplier
PhytoComm	Mao Dong Qing	G319HS063SK1	62620	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Mao Dong Qing*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Mao Dong Qing*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mao Dong Qing	G319HS063RN1	62417†	20
PhytoComm	Mao Dong Qing	G319HS063RN1	62418†	20
PhytoComm	Mao Dong Qing	G319HS063SK1	62619†	20
PhytoComm	Mao Dong Qing	G319HS063SK1	62620†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Mao Dong Qing*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mao Dong Qing* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mao Dong Qing* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	3	79	1	12 394
Type C	0	0	0	857

The substance/substance group *Mao Dong Qing* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9667 % (> 99.9368 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mao Dong Qing* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Du Zhong	6.78	–
Bai Xian Pi	7.07	–
Yin Yang Huo	7.84	–
Ling Zhi	7.88	–
Sha Ren	9.15	–
Ye Jiao Teng	9.26	–
Yu Jin	9.45	–
Che Qian Zi	9.98	–
Hou Po	10.95	–
Qing Hao	11.72	–
Qiang Huo	12.31	–
Bo He	12.36	–
Dan Dou Chi	12.67	–
Chai Hu	13.08	–
Zi Hua Di Ding	14.31	–
Jiao Gu Lan	15.07	–
Jing Jie	15.45	–
Fu Zi	16.22	–
Shen Qu	16.33	–
Ren Dong Teng	16.95	–
Tu Fu Ling	17.56	–
Ce Bai Ye	17.58	–
Xi Xian Cao	17.62	–
Qing Pi	17.74	–
Yan Hu Suo	18.13	–
Bai Jiang Cao	18.42	–
Lian Zi	18.54	–
Ji Li	18.71	–
Huang Bai	18.82	–
Shan Yao	19.83	–
Zhu Ling	20.14	–
Ji Xue Teng	20.33	–
Rou Gui	20.90	–
Dan Shen	21.23	–
Guang Huo Xiang	21.90	–
Pi Pa Ye	22.05	–
Tian Hua Fen	23.07	–
Zhi Ke	23.11	–
Hong Jing Tian	23.13	–
Nü Zhen Zi	23.27	–
He Huan Pi	23.52	–
Fu Pen Zi	23.68	–
Ma Huang	23.79	–
Huang Lian	23.95	–
Pu Gong Ying	24.65	–
Chuan Xiong	25.26	–
Wu Mei	25.42	–
Gou Teng	25.58	–
Lian Qiao	26.13	–
Wu Wei Zi	26.25	–
Ze Lan	26.32	–
Ma Huang Gen	26.67	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Di Gu Pi	26.75	–
Jin Yin Hua	27.85	–
(Fen) Bi Xie	28.14	–
Gu Sui Bu	28.65	–
Sang Ye	29.04	–
Suan Zao Ren	29.28	–
Sheng Jiang	30.07	–
Yi Yi Ren	30.25	–
Bai Shao Yao	30.97	–
Ban Lan Gen	31.00	–
Mang Xiao	31.97	–
Ban Zhi Lian	32.04	–
Zhi Shi	32.10	–
Lu Gen	32.35	–
Ren Shen	32.58	–
Cang Er Zi	33.23	–
Gui Zhi	33.32	–
Yuan Zhi	33.37	–
Huo Ma Ren	33.53	–
Chuang Mu Xiang	34.08	–
Gua Lou	34.20	–
Gan Cao	34.83	–
Xin Yi	35.10	–
Mu Zei	35.31	–
Zhe Bei Mu	35.31	–
E Zhu	35.71	–
Cang Zhu	35.99	–
Fu Ling	36.53	–
Gou Qi Zi	37.01	–
Jiang Huang	37.27	–
She Gan	38.42	–
Tai Zi Shen	38.74	–
Tao Ren	39.41	–
Yin Chen Hao	39.46	–
Fu Shen	40.29	–
Hong Hua	40.78	–
Wang Bu Liu Xing	41.34	–
Chi Shao (Yao)	43.63	–
Shan Yu Rou	43.87	–
Sang Zhi	44.00	–
Zhi Gan Cao	44.12	–
Niu Bang Zi	44.66	–
Fo Shou	45.29	–
Ban Xia (Jiang)	45.95	–
Ci Wu Jia	45.99	–
Yi Mu Cao	46.29	–
Fu Xiao Mai	46.40	–
Xie Bai	46.94	–
Jie Geng	47.81	–
Zi Su Zi	48.07	–
Chen Pi	48.38	–
Yu Xing Cao	49.04	–
Huang Qin	49.06	–
Zhu Ru	49.31	–
Bai Zi Ren	49.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mao Dong Qing* is separated from critical neighbours in a second-stage model,

all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62417	62417	0.00	6.78
62418	62418	0.00	7.07
62619	62619	0.00	9.89
62620	62620	0.00	9.89

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mi Huan Jun**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60556-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mi Huan Jun; Armillariella mellea rhizoma

Special notes

When selecting the *Mi Huan Jun* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mi Huan Jun	1	0	2

Second-stage model

For differentiation of the substance/substance group *Mi Huan Jun* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mi Huan Jun*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mi Huan Jun	G510H0418021	62983	40	from supplier
PhytoComm	Mi Huan Jun	G510H0418021	62984	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Mi Huan Jun*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Mi Huan Jun*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mi Huan Jun	G510H0418021	62983†	20
PhytoComm	Mi Huan Jun	G510H0418021	62984†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Mi Huan Jun*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mi Huan Jun	G510H0418221	1
Phytocomm	Mi Huan Jun	G510H0418422	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mi Huan Jun* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mi Huan Jun* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	79	1	24 520
Type B	0	40	0	12 437
Type C	1	0	2	854

The substance/substance group *Mi Huan Jun* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.9070 % (> 99.3221 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mi Huan Jun* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chuan Mu Tong	5.98	–
Sang Zhi	9.05	–
Di Gu Pi	10.01	–
Chuan Lian Zi	11.52	–
Mai Ya	13.11	–
E Zhu	14.60	–
Bai He	15.01	–
Gan Jiang	15.31	–
Dang Gui	16.59	–
Ji Li	17.99	–
Sha Shen (Bei)	18.14	–
Long Dan (Cao)	20.77	–
Qin Jiao	20.98	–
Tian Hua Fen	21.86	–
Lai Fu Zi	22.58	–
Ban Zhi Lian	23.66	–
Mang Xiao	24.29	–
Yan Hu Suo	24.29	–
(Huai) Niu Xi	25.54	–
Zi Su Zi	27.20	–
Jiang Huang	28.18	–
Cang Er Zi	28.54	–
Huang Qi	30.21	–
Zhi Gan Cao	30.26	–
Da Zao	30.29	–
Shan Yao	30.67	–
Chuan Niu Xi	30.76	–
Jie Geng	31.35	–
Bai Zhi	31.69	–
(Shi) Chang Pu	31.85	–
Bai Zhu	32.36	–
(Bai) Dou Kou	33.52	–
Niu Bang Zi	35.51	–
Fu Ling	36.34	–
Sha Ren	36.35	–
Chen Pi	36.83	–
He Huan Pi	37.71	–
Lian Qiao	38.17	–
Lian Zi	38.63	–
Rou Gui	39.05	–
Gua Lou	41.72	–
Bai Xian Pi	41.98	–
Bing Lang	42.26	–
Suan Zao Ren	42.94	–
Yi Yi Ren	43.08	–
Bai Shao Yao	43.74	–
Xiao Hui Xiang	44.41	–
Wu Wei Zi	45.25	–
Xiang Fu	45.98	–
Tu Fu Ling	46.39	–
Zi Hua Di Ding	46.59	–
Chi Shao (Yao)	46.83	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Lian	47.57	–
Mu Dan Pi	47.61	–
Ku Shen	47.63	–
Jin Yin Hua	47.74	–
Ye Jiao Teng	47.99	–
Mu Gua	48.68	–
Sheng Jiang	48.72	–
Yuan Zhi	48.83	–
Ji Xue Teng	48.86	–
Bai Hua She She Cao	48.91	–
Du Zhong	48.93	–
Mao Dong Qing	49.12	–
Qiang Huo	49.17	–
Zhi Ke	49.62	–
Chai Hu	49.66	–
Bo He	49.72	–
Shan Yu Rou	50.06	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mi Huan Jun* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62983	62983	0.00	6.14
62984	62984	0.00	5.98

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mu Dan Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60113-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mu Dan Pi; Moutan cortex radicis

Special notes

When selecting the *Mu Dan Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mu Dan Pi	1	0	5

Second-stage model

For differentiation of the substance/substance group *Mu Dan Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mu Dan Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mu Dan Pi	G166H0741021	62825	40	from supplier
PhytoComm	Mu Dan Pi	G166H0741021	62826	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Mu Dan Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Mu Dan Pi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mu Dan Pi	G166H0741021	62825†	20
PhytoComm	Mu Dan Pi	G166H0741021	62826†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 11 spectra from 4 *Apo-Ident* customers from 5 batches from the substance/substance group *Mu Dan Pi*.
- Among them are spectra of independent samples from 5 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mu Dan Pi	G166H0741121	1
Phytocomm	Mu Dan Pi	G166H0741122	2
Phytocomm	Mu Dan Pi	G166H0741321	2
PhytoComm	Mu Dan Pi	G166H0741321	1
Phytocomm	Mu Dan Pi	G166H0741421	4
Phytocomm	Mu Dan Pi	g166j074122	1

- 846 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mu Dan Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mu Dan Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	11	846

The substance/substance group *Mu Dan Pi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8249 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mu Dan Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chen Pi	11.55	–
Shan Yu Rou	17.57	–
Gu Sui Bu	18.71	–
He Huan Pi	20.14	–
Wu Yao	20.66	–
He Shou Wu	21.36	–
Dang Gui	22.82	–
(Shi) Chang Pu	23.78	–
Ju Hua	24.56	–
Chuan Lian Zi	24.87	–
Ji Li	26.15	–
Mang Xiao	26.44	–
Fang Feng	26.46	–
Hong Jing Tian	26.66	–
Ba Ji Tian	26.71	–
Sang Ji Shend	27.44	–
Sha Ren	27.46	–
Jiang Huang	27.76	–
Huang Qi	31.64	–
Chuan Niu Xi	32.14	–
Bing Lang	33.00	–
E Zhu	34.46	–
Sang Zhi	34.64	–
Du Huo	34.80	–
Yin Chen Hao	34.85	–
Xiang Fu	34.91	–
Wu Zhu Yu	35.44	–
Tu Si Zi	36.12	–
Shan Yao	36.78	–
Sang Bai Pi	38.63	–
(Bai) Dou Kou	40.38	–
Ge Gen	40.48	–
Zhi Gan Cao	40.81	–
Gua Lou	41.24	–
Tian Hua Fen	42.15	–
Zhi Ke	42.94	–
Bai Hua She She Cao	43.03	–
Yan Hu Suo	43.39	–
Xiao Hui Xiang	44.01	–
Ku Shen	44.58	–
Di Gu Pi	45.13	–
(Huai) Niu Xi	45.18	–
Lai Fu Zi	45.28	–
Yi Mu Cao	45.31	–
Da Zao	45.45	–
Zi Su Zi	45.94	–
Bai Zhi	46.20	–
Xin Yi	47.71	–
Xu Duan	47.99	–
Dan Zhu Ye	49.74	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mu Dan Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62825	62825	0.00	11.55
62826	62826	0.00	11.74

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mu Gua**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10003578-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mu Gua; Chaenomelis lagenariae fructus

Special notes

When selecting the *Mu Gua* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mu Gua	2	0	2

Second-stage model

For differentiation of the substance/substance group *Mu Gua* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mu Gua*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mu Gua	G063HS043TK1	62929	40	from supplier
PhytoComm	Mu Gua	G063HS043TK1	62930	40	from supplier
PhytoComm	Mu Gua	G063H0424922	63017	40	from supplier
PhytoComm	Mu Gua	G063H0424922	63018	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Mu Gua*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Mu Gua*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mu Gua	G063HS043TK1	62929†	20
PhytoComm	Mu Gua	G063HS043TK1	62930†	20
PhytoComm	Mu Gua	G063H0424922	63017†	20
PhytoComm	Mu Gua	G063H0424922	63018†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Mu Gua*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mu Gua	G063H0424221	1
PhytoComm	Mu Gua	G063H0424221	1
Phytocomm	Mu Gua	G063H0424521	4

- 851 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mu Gua* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mu Gua* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	6	851

The substance/substance group *Mu Gua* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mu Gua* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Huang Qi	7.73	–
Shan Yu Rou	8.76	–
Cang Zhu	10.97	–
Zhi Mu	11.28	–
Bai Shao Yao	11.66	–
Zhi Gan Cao	12.08	–
Bai Zhu	12.92	–
Gou Qi Zi	13.37	–
Dang Gui	13.92	–
San Qi	14.11	–
Yuan Zhi	14.27	–
Chi Shao (Yao)	14.34	–
Ren Shen	14.52	–
Jie Geng	14.90	–
Pi Pa Ye	15.71	–
Xie Bai	15.97	–
Dang Gui Wei	16.01	–
Chuang Mu Xiang	16.26	–
Shan Yao	17.79	–
Chen Pi	18.40	–
Ku Shen	18.46	–
Chuan Xiong	18.49	–
Gua Lou	19.40	–
Lian Qiao	19.42	–
Yu Zhu	20.22	–
Chuan Lian Zi	20.48	–
Sang Zhi	21.56	–
Jiao Gu Lan	21.64	–
Zhe Bei Mu	21.78	–
Di Gu Pi	22.53	–
Bai Zi Ren	22.71	–
Ji Li	23.08	–
Suan Zao Ren	23.11	–
Fo Shou	23.65	–
(Huai) Niu Xi	24.32	–
Lian Zi	24.42	–
Gan Cao	24.61	–
Mu Zei	24.63	–
Ye Jiao Teng	24.76	–
Ban Lan Gen	24.92	–
Wu Wei Zi	25.22	–
Ren Dong Teng	25.41	–
Long Dan (Cao)	25.44	–
Jin Yin Hua	25.64	–
Tian Hua Fen	25.86	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	26.04	–
Zhu Ru	26.09	–
Ling Zhi	26.10	–
Long Yan Rou	26.20	–
Tai Zi Shen	26.20	–
Zhi Ke	26.52	–
Lai Fu Zi	26.99	–
Shen Qu	27.02	–
He Huan Pi	27.23	–
Qin Jiao	27.60	–
Fu Zi	27.84	–
Dan Dou Chi	27.96	–
Mao Dong Qing	28.23	–
Mang Xiao	28.52	–
Fu Ling	29.12	–
Dan Shen	29.28	–
Hou Po	29.60	–
Sheng Jiang	29.66	–
Huo Ma Ren	29.71	–
Bai He	30.41	–
Ze Xie	31.47	–
Lu Gen	32.30	–
Fu Pen Zi	32.46	–
Ci Wu Jia	32.66	–
Tao Ren	33.37	–
Chai Hu	33.63	–
Huang Qin	33.85	–
Hong Jing Tian	34.10	–
Rou Gui	34.55	–
Fu Xiao Mai	34.95	–
Bing Lang	35.00	–
Mai Men Dong	35.31	–
Ban Xia (Jiang)	36.26	–
Ji Xue Teng	36.71	–
Gou Teng	36.91	–
Zi Su Zi	37.35	–
Che Qian Zi	37.64	–
Gui Zhi	38.31	–
Bai Xian Pi	38.80	–
Chuan Mu Tong	38.96	–
Yin Yang Huo	39.94	–
Sha Shen (Bei)	40.08	–
Yan Hu Suo	40.52	–
Hong Hua	40.59	–
Yi Yi Ren	40.70	–
Ma Huang	40.95	–
Gan Jiang	41.37	–
Gu Sui Bu	41.38	–
Qing Pi	41.45	–
Cang Er Zi	41.83	–
Qiang Huo	41.86	–
Zi Hua Di Ding	42.16	–
Ban Zhi Lian	42.20	–
Bai Zhi	42.48	–
She Gan	42.75	–
Ce Bai Ye	43.27	–
Chuan Niu Xi	43.52	–
(Shi) Chang Pu	44.10	–
E Zhu	44.35	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
(Fen) Bi Xie	44.81	–
Yu Jin	45.39	–
Bo He	45.43	–
Zhu Ling	46.17	–
Guang Huo Xiang	48.42	–
Huang Bai	48.51	–
Huang Lian	48.63	–
Ma Huang Gen	49.00	–
Jing Jie	50.59	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mu Gua* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62929	62929	0.00	9.48
62930	62930	0.00	8.76
63017	63017	0.00	8.30
63018	63018	0.00	7.73

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Mu Zei**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60123-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Mu Zei; Equiseti hiemalis herba

Special notes

When selecting the *Mu Zei* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Mu Zei	1	0	2

Second-stage model

For differentiation of the substance/substance group *Mu Zei* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Mu Zei*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mu Zei	G265HS045TK1	62811	40	from supplier
PhytoComm	Mu Zei	G265HS045TK1	62812	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Mu Zei*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Mu Zei*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Mu Zei	G265HS045TK1	62811†	20
PhytoComm	Mu Zei	G265HS045TK1	62812†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Mu Zei*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Mu Zei	G265H0427421	1
PhytoComm	Mu Zei	G265H0427823	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Mu Zei* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Mu Zei* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	76	4	24 517
Type B	2	18	22	12 435
Type C	1	0	2	854

The substance/substance group *Mu Zei* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9911 % (> 99.9761 %)	95.0000 % (> 91.2500 %)
Type B	99.9714 % (> 99.9417 %)	45.0000 % (> 37.5000 %)
Type C	99.9070 % (> 99.3221 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Mu Zei* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shen Qu	3.95	–
Huo Ma Ren	4.29	–
He Huan Pi	4.60	–
Bai Shao Yao	5.15	–
Tai Zi Shen	5.28	–
Lai Fu Zi	6.30	–
Lu Gen	6.47	–
Tao Ren	6.72	–
Gui Zhi	7.19	–
Ye Jiao Teng	7.22	–
Gou Teng	7.54	–
Zhe Bei Mu	7.60	–
Gu Sui Bu	7.61	–
Ji Li	8.16	–
Ji Xue Teng	8.37	–
Tian Hua Fen	8.72	–
Ci Wu Jia	8.88	–
Sheng Jiang	9.21	–
Bai Zi Ren	9.85	–
Ling Zhi	9.98	–
Rou Gui	10.91	–
Lian Zi	10.95	–
Ze Xie	10.99	–
Fo Shou	11.27	–
Gua Lou	11.39	–
Fu Ling	11.58	–
Ren Dong Teng	11.58	–
Ban Xia (Jiang)	11.83	–
Di Gu Pi	11.99	–
Bai Xian Pi	12.63	–
(Fen) Bi Xie	12.64	–
Lian Qiao	13.35	–
Pi Pa Ye	13.39	–
Zhu Ru	13.83	–
Shan Yao	14.01	–
Ma Huang Gen	14.26	–
Chuan Xiong	14.49	–
Suan Zao Ren	14.54	–
Fu Xiao Mai	16.02	–
She Gan	16.19	–
Yi Yi Ren	16.57	–
Jin Yin Hua	16.87	–
Chen Pi	16.90	–
Tu Fu Ling	16.97	–
Fu Zi	17.27	–
Zhu Ling	17.56	–
Ce Bai Ye	17.80	–
Yuan Zhi	18.12	–
Fu Pen Zi	18.21	–
Yan Hu Suo	18.27	–
Yu Jin	18.28	–
Dan Shen	19.55	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Dou Chi	20.13	–
Ban Lan Gen	20.69	–
Yu Zhu	21.21	–
Dang Gui Wei	21.38	–
Yin Yang Huo	21.78	–
Jiao Gu Lan	21.98	–
Ma Huang	22.31	–
Hou Po	22.64	–
Jie Geng	23.08	–
Cang Zhu	23.45	–
Hong Jing Tian	23.69	–
Chai Hu	23.89	–
Dang Gui	24.00	–
Guang Huo Xiang	24.08	–
Mao Dong Qing	24.74	–
Ren Shen	24.89	–
Fu Shen	24.89	–
Che Qian Zi	25.26	–
Zhi Ke	25.68	–
San Qi	25.80	–
Zhi Mu	27.75	–
Gan Cao	28.42	–
Qiang Huo	29.54	–
Huang Qin	29.68	–
Wu Wei Zi	30.00	–
Shan Yu Rou	31.29	–
Mang Xiao	32.21	–
Gou Qi Zi	32.59	–
Mu Gua	33.33	–
Jing Jie	33.80	–
Ku Shen	33.98	–
Long Yan Rou	34.26	–
Zi Hua Di Ding	36.38	–
Huang Lian	37.11	–
Sang Zhi	37.98	–
Chuang Mu Xiang	37.99	–
Sha Ren	38.79	–
Qing Pi	39.24	–
Zhi Gan Cao	39.25	–
Xie Bai	39.66	–
Chi Shao (Yao)	42.65	–
Du Zhong	42.92	–
Bo He	42.99	–
Huang Bai	48.01	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mu Zei* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62811	62811	0.00	4.74
62812	62812	0.00	3.95

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Niu Bang Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002261-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Niu Bang Zi; Arctii fructus

Special notes

When selecting the *Niu Bang Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Niu Bang Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Niu Bang Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Niu Bang Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Niu Bang Zi	G026H0433822	62765	40	from supplier
PhytoComm	Niu Bang Zi	G026H0433822	62766	40	from supplier
PhytoComm	Niu Bang Zi	G026H0433921	62967	40	from supplier
PhytoComm	Niu Bang Zi	G026H0433921	62968	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Niu Bang Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Niu Bang Zi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Niu Bang Zi	G026H0433822	62765†	20
PhytoComm	Niu Bang Zi	G026H0433822	62766†	20
PhytoComm	Niu Bang Zi	G026H0433921	62967†	20
PhytoComm	Niu Bang Zi	G026H0433921	62968†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Niu Bang Zi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
phytocomm	Niu Bang Zi	g026h0433321	2
Phytocomm	Niu Bang Zi	G026H0433321	1
PhytoComm	Niu Bang Zi	G026H0433321	2
Phytocomm	Niu Bang Zi	G026H0433521	3

- 849 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Niu Bang Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Niu Bang Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	2	6	849

The substance/substance group *Niu Bang Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8253 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Niu Bang Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lian Qiao	11.55	–
Ku Shen	12.63	–
Gua Lou	12.92	–
Zi Su Zi	18.16	–
Cang Er Zi	19.19	–
Yi Mu Cao	20.14	–
Xiao Hui Xiang	21.01	–
Shan Yao	24.35	–
Di Gu Pi	24.79	–
Xiang Fu	25.44	–
(Huai) Niu Xi	27.20	–
Jiang Huang	30.28	–
Yan Hu Suo	30.45	–
Mang Xiao	30.94	–
Gan Jiang	31.69	–
Gan Cao	31.97	–
Tian Hua Fen	32.04	–
Sang Zhi	33.62	–
Zhi Ke	33.63	–
Wu Wei Zi	34.29	–
Dong Gua Zi	34.83	–
Wu Mei	35.29	–
Chuan Lian Zi	35.68	–
(Bai) Dou Kou	36.18	–
Jie Geng	36.26	–
Suan Zao Ren	36.40	–
Sha Ren	36.57	–
Shen Qu	36.74	–
Jin Yin Hua	36.87	–
Pu Gong Ying	37.57	–
Chi Shao (Yao)	37.63	–
E Zhu	37.96	–
(Shi) Chang Pu	38.35	–
Ji Li	38.86	–
Qiang Huo	38.98	–
Huang Qi	39.59	–
Yuan Zhi	39.69	–
Ze Lan	40.43	–
Chen Pi	40.50	–
Zi Hua Di Ding	41.12	–
Zhi Gan Cao	42.56	–
Shan Yu Rou	42.68	–
Chuan Xiong	43.00	–
Huang Lian	43.01	–
Du Huo	43.54	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Ye	43.61	–
Ju Hua	43.93	–
Qing Pi	44.13	–
Bai Zhi	44.32	–
Wu Yao	44.75	–
Sha Shen (Bei)	45.42	–
Dang Gui	46.06	–
Huang Bai	46.60	–
Chuan Niu Xi	47.77	–
Chuan Mu Tong	48.78	–
Jiao Gu Lan	49.05	–
Bai Hua She She Cao	49.07	–
Fu Zi	49.13	–
Xie Bai	49.19	–
Du Zhong	49.25	–
Xu Duan	49.54	–
Wang Bu Liu Xing	49.79	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Niu Bang Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62765	62765	0.00	11.55
62766	62766	0.00	11.77
62967	62967	0.00	12.63
62968	62968	0.00	12.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Nü Zhen Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60065-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Nü Zhen Zi; Ligustri lucidi fructus

Special notes

When selecting the *Nü Zhen Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Nü Zhen Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Nü Zhen Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Nü Zhen Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Nü Zhen Zi	G138H0345821	62343	40	from supplier
PhytoComm	Nü Zhen Zi	G138H0345821	62344	40	from supplier
PhytoComm	Nü Zhen Zi	G138HS010SK1	62569	40	from supplier
PhytoComm	Nü Zhen Zi	G138HS010SK1	62570	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Nü Zhen Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Nü Zhen Zi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Nü Zhen Zi	G138H0345821	62343†	20
PhytoComm	Nü Zhen Zi	G138H0345821	62344†	20
PhytoComm	Nü Zhen Zi	G138HS010SK1	62569†	20
PhytoComm	Nü Zhen Zi	G138HS010SK1	62570†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Nü Zhen Zi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Nü Zhen Zi	420319901	1
Phytocomm	Nü Zhen Zi	G138H0345421	1
PhytoComm	Nü Zhen Zi	G138H0345421	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Nü Zhen Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Nü Zhen Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	3	854

The substance/substance group *Nü Zhen Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Nü Zhen Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shan Yao	16.51	–
Yin Chen Hao	16.99	–
Xi Xian Cao	17.29	–
Sang Ye	17.70	–
Yi Mu Cao	18.35	–
Du Zhong	19.38	–
Xin Yi	20.42	–
Zhi Shi	20.82	–
Sang Bai Pi	21.08	–
Wu Mei	22.00	–
Qiang Huo	22.32	–
Pu Gong Ying	23.51	–
Jing Jie	23.62	–
Qing Pi	23.65	–
Mao Dong Qing	24.14	–
Sha Ren	24.22	–
Wu Zhu Yu	24.37	–
Gou Teng	24.40	–
Dan Zhu Ye	24.66	–
Zi Hua Di Ding	25.03	–
Mang Xiao	26.07	–
Jiang Huang	26.58	–
Ce Bai Ye	27.35	–
Huang Lian	28.06	–
Ze Lan	28.26	–
Tu Fu Ling	28.41	–
Dan Shen	28.51	–
Yu Jin	28.90	–
Hou Po	29.21	–
Sang Ji Shend	29.24	–
Fu Zi	29.67	–
Bai Jiang Cao	29.71	–
Dan Dou Chi	29.71	–
Chai Hu	30.11	–
Zhi Ke	30.69	–
Wu Wei Zi	31.23	–
Qing Hao	31.31	–
Hong Jing Tian	31.54	–
Cang Er Zi	31.77	–
Xiao Hui Xiang	31.81	–
He Huan Pi	31.90	–
Hu Zhang	31.90	–
Huang Bai	32.35	–
Jiao Gu Lan	32.70	–
Hong Hua	32.76	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xuan Fu Hua	32.82	–
Ling Zhi	32.96	–
Bo He	33.05	–
Yu Xing Cao	33.07	–
Jin Yin Hua	33.77	–
Ye Jiao Teng	34.56	–
Yin Yang Huo	34.87	–
Di Gu Pi	35.03	–
Wang Bu Liu Xing	35.17	–
Bai Hua She She Cao	36.85	–
Bu Gu Zhi	37.64	–
Fu Pen Zi	37.91	–
Han Lian Cao	38.37	–
Gu Sui Bu	38.39	–
Ma Huang	38.80	–
Bai Xian Pi	39.28	–
Zi Su Zi	39.98	–
He Shou Wu	40.09	–
Lian Qiao	40.37	–
Xia Ku Cao	40.64	–
Ku Shen	41.14	–
Chuang Mu Xiang	41.80	–
Chuan Xiong	42.05	–
Zhe Bei Mu	42.79	–
Shan Yu Rou	43.55	–
Che Qian Zi	43.71	–
Gua Lou	43.81	–
Chen Pi	44.24	–
Ban Lan Gen	44.35	–
Pi Pa Ye	44.63	–
Yan Hu Suo	45.33	–
(Shi) Chang Pu	46.19	–
Ge Gen	46.48	–
Wu Jia Pi	46.62	–
Rou Cong Rong	46.77	–
Ji Li	47.00	–
Ren Dong Teng	47.80	–
E Zhu	47.85	–
Ding Xiang	48.14	–
Niu Bang Zi	48.20	–
Shen Qu	48.41	–
Bai Shao Yao	48.54	–
Jin Qian Cao	48.54	–
Tian Hua Fen	48.60	–
Lian Zi	48.73	–
Guang Huo Xiang	48.76	–
Fang Feng	49.28	–
Gan Cao	49.43	–
Jie Geng	49.57	–
Cang Zhu	49.77	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Nü Zhen Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62343	62343	0.00	18.35
62344	62344	0.00	18.53
62569	62569	0.00	16.65
62570	62570	0.00	16.51

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Pi Pa Ye**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60201-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Pi Pa Ye; Eriobotryae japonicae folium

Special notes

When selecting the *Pi Pa Ye* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Pi Pa Ye	1	0	1

Second-stage model

For differentiation of the substance/substance group *Pi Pa Ye* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Pi Pa Ye*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Pi Pa Ye	G099HS166RP1	62451	40	from supplier
PhytoComm	Pi Pa Ye	G099HS166RP1	62452	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Pi Pa Ye*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Pi Pa Ye*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Pi Pa Ye	G099HS166RP1	62451†	20
PhytoComm	Pi Pa Ye	G099HS166RP1	62452†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Pi Pa Ye*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Pi Pa Ye	G099H0832221	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Pi Pa Ye* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pi Pa Ye* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	4	78	2	24 516
Type B	6	32	8	12 431
Type C	0	0	1	856

The substance/substance group *Pi Pa Ye* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9833 % (> 99.9684 %)	97.5000 % (> 93.7500 %)
Type B	99.9310 % (> 99.9012 %)	80.0000 % (> 72.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pi Pa Ye* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Shan Yao	3.98	–
Lian Qiao	5.75	–
Ze Xie	6.39	–
Fu Zi	6.80	–
Tian Hua Fen	7.62	–
Ji Li	7.98	–
Ling Zhi	8.26	–
Ren Dong Teng	8.40	–
Zhe Bei Mu	8.55	–
Gua Lou	8.73	–
Chuan Xiong	8.90	–
Bai Zi Ren	9.45	–
Mu Zei	9.48	–
Shen Qu	9.82	–
Huo Ma Ren	10.19	–
Cang Zhu	10.45	–
Chen Pi	11.03	–
Bai Shao Yao	11.09	–
Yuan Zhi	11.48	–
Ban Lan Gen	11.87	–
Jin Yin Hua	12.13	–
San Qi	12.65	–
Lai Fu Zi	12.79	–
Hou Po	13.05	–
Lian Zi	13.32	–
He Huan Pi	13.42	–
Tai Zi Shen	13.53	–
Jiao Gu Lan	13.68	–
Sheng Jiang	14.21	–
Fu Ling	14.53	–
Dan Dou Chi	14.58	–
Gu Sui Bu	14.70	–
Fu Pen Zi	15.07	–
Dan Shen	16.13	–
Ku Shen	16.21	–
Zhu Ru	16.34	–
Suan Zao Ren	16.43	–
Dang Gui	16.67	–
Yu Zhu	17.02	–
Mao Dong Qing	17.22	–
Ye Jiao Teng	17.30	–
Ren Shen	17.36	–
Dang Gui Wei	17.42	–
Chai Hu	17.66	–
Che Qian Zi	17.70	–
Di Gu Pi	18.17	–
Gou Qi Zi	18.19	–
Jie Geng	18.25	–
Tu Fu Ling	18.25	–
Ji Xue Teng	18.30	–
Ci Wu Jia	18.62	–
Gui Zhi	18.62	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lu Gen	18.64	–
Hong Jing Tian	19.28	–
Gou Teng	19.69	–
Yan Hu Suo	19.83	–
Wu Wei Zi	19.85	–
Tao Ren	19.87	–
Yu Jin	20.07	–
Ce Bai Ye	20.12	–
Zhi Gan Cao	20.51	–
Xie Bai	20.83	–
She Gan	21.77	–
Zhi Mu	21.89	–
Bai Xian Pi	22.08	–
Zhu Ling	22.16	–
Guang Huo Xiang	22.28	–
Yin Yang Huo	22.44	–
Fu Xiao Mai	23.03	–
Qiang Huo	23.61	–
Yi Yi Ren	23.85	–
Zhi Ke	24.26	–
Shan Yu Rou	24.28	–
Ban Xia (Jiang)	24.34	–
(Fen) Bi Xie	24.41	–
Fo Shou	24.42	–
Rou Gui	24.98	–
Ma Huang	24.99	–
Gan Cao	25.39	–
Mu Gua	26.17	–
Chuang Mu Xiang	26.53	–
Qing Pi	26.77	–
Long Yan Rou	27.93	–
Huang Qin	28.35	–
Ma Huang Gen	29.10	–
Zi Hua Di Ding	30.22	–
Mang Xiao	32.73	–
Bo He	32.80	–
Jing Jie	33.32	–
Sang Zhi	33.50	–
Huang Lian	33.73	–
Du Zhong	34.80	–
Chi Shao (Yao)	35.39	–
Sha Ren	35.98	–
Huang Bai	37.85	–
Cang Er Zi	38.88	–
Hong Hua	39.35	–
Fu Shen	42.14	–
Sang Ye	47.50	–
Bai Jiang Cao	48.56	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pi Pa Ye* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62451	62451	0.00	4.23
62452	62452	0.00	3.98

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Pu Gong Ying**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50358-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Pu Gong Ying; Taraxaci mongolici herba cum radice

Special notes

When selecting the *Pu Gong Ying* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Pu Gong Ying	2	0	1

Second-stage model

For differentiation of the substance/substance group *Pu Gong Ying* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Pu Gong Ying*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Pu Gong Ying	G239HS308RN1	62263	40	from supplier
PhytoComm	Pu Gong Ying	G239HS308RN1	62264	40	from supplier
PhytoComm	Pu Gong Ying	G239HS308SH1	62635	40	from supplier
PhytoComm	Pu Gong Ying	G239HS308SH1	62636	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Pu Gong Ying*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Pu Gong Ying*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Pu Gong Ying	G239HS308RN1	62263†	20
PhytoComm	Pu Gong Ying	G239HS308RN1	62264†	20
PhytoComm	Pu Gong Ying	G239HS308SH1	62635†	20
PhytoComm	Pu Gong Ying	G239HS308SH1	62636†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 1 batches from the substance/substance group *Pu Gong Ying*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Pu Gong Ying	G239H1315321	5
PhytoComm	Pu Gong Ying	G239H1315321	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Pu Gong Ying* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Pu Gong Ying* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	6	851

The substance/substance group *Pu Gong Ying* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Pu Gong Ying* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yin Chen Hao	7.31	–
Sang Ye	7.51	–
Ze Lan	8.02	–
Huang Lian	8.23	–
Xin Yi	9.08	–
Tu Fu Ling	9.39	–
Xi Xian Cao	11.30	–
Nü Zhen Zi	14.19	–
Shan Yao	14.39	–
Bai Jiang Cao	14.82	–
Zhi Shi	15.89	–
Sha Ren	16.78	–
Dan Shen	18.78	–
Han Lian Cao	19.85	–
Hu Zhang	20.59	–
Bo He	20.95	–
Jiang Huang	21.10	–
Zhi Ke	21.25	–
Fu Zi	21.42	–
Yu Xing Cao	21.85	–
Du Zhong	22.34	–
Huang Bai	23.75	–
Che Qian Zi	23.91	–
Qiang Huo	24.05	–
Bai Hua She She Cao	24.27	–
Jiao Gu Lan	25.53	–
Mao Dong Qing	25.64	–
Yu Jin	25.65	–
Mang Xiao	28.54	–
Chai Hu	28.68	–
Xiang Fu	29.49	–
Hou Po	29.69	–
Wu Mei	29.88	–
Wu Wei Zi	29.97	–
Qing Hao	30.32	–
Yi Mu Cao	31.51	–
Hong Jing Tian	31.51	–
Zi Hua Di Ding	31.84	–
Qing Pi	32.32	–
Ye Jiao Teng	32.39	–
Yin Yang Huo	32.47	–
Jing Jie	33.19	–
Dan Dou Chi	33.72	–
Cang Er Zi	34.21	–
Fu Pen Zi	35.78	–
Chuang Mu Xiang	36.68	–
Jin Yin Hua	36.71	–
Ma Huang	36.72	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yan Hu Suo	36.81	–
Zi Su Zi	36.93	–
Gua Lou	37.25	–
Di Gu Pi	37.82	–
Bai Xian Pi	38.38	–
(Shi) Chang Pu	38.53	–
Suan Zao Ren	38.56	–
Ce Bai Ye	38.89	–
Wang Bu Liu Xing	39.24	–
Xia Ku Cao	39.34	–
He Huan Pi	40.27	–
Chi Shao (Yao)	40.99	–
Niu Bang Zi	41.76	–
E Zhu	41.93	–
Chuan Xiong	42.76	–
Pi Pa Ye	43.00	–
Gu Sui Bu	43.10	–
Gan Cao	43.13	–
Hong Hua	43.64	–
Lian Zi	43.99	–
Bai Shao Yao	45.79	–
Sang Ji Shend	45.84	–
Lian Qiao	46.18	–
(Bai) Dou Kou	46.54	–
Gou Qi Zi	47.32	–
Ren Dong Teng	47.48	–
Ling Zhi	47.62	–
Huang Qin	47.81	–
Shan Yu Rou	48.46	–
Ji Li	48.97	–
Tian Hua Fen	49.09	–
Cang Zhu	49.88	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Pu Gong Ying* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62263	62263	0.00	8.13
62264	62264	0.00	8.02
62635	62635	0.00	7.31
62636	62636	0.00	7.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by

laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Qiang Huo**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60109-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Qiang Huo; Notopterygii rhizoma et radix

Special notes

When selecting the *Qiang Huo* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Qiang Huo	1	0	2

Second-stage model

For differentiation of the substance/substance group *Qiang Huo* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Qiang Huo*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Qiang Huo	G172HS154SL1	62719	40	from supplier
PhytoComm	Qiang Huo	G172HS154SL1	62720	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Qiang Huo*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Qiang Huo*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Qiang Huo	G172HS154SL1	62719†	20
PhytoComm	Qiang Huo	G172HS154SL1	62720†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Qiang Huo*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Qiang Huo	G172H0840321	1
PhytoComm	Qiang Huo	G172H0840321	1
Phytocomm	Qiang Huo	g172h0841321	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Qiang Huo* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Qiang Huo* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	1	38	2	12 436
Type C	0	0	3	854

The substance/substance group *Qiang Huo* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9929 % (> 99.9631 %)	95.0000 % (> 87.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Qiang Huo* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mao Dong Qing	6.07	–
Zi Hua Di Ding	6.94	–
Hou Po	7.50	–
Jing Jie	7.67	–
Chai Hu	7.81	–
Yu Jin	7.90	–
Ling Zhi	8.24	–
Bo He	8.43	–
Du Zhong	8.49	–
Qing Hao	9.72	–
Yin Yang Huo	10.05	–
Bai Xian Pi	10.19	–
Dan Dou Chi	10.35	–
Ce Bai Ye	11.28	–
Qing Pi	11.71	–
Fu Zi	11.83	–
Sha Ren	13.31	–
Ye Jiao Teng	13.75	–
Dan Shen	14.85	–
Shen Qu	15.55	–
Lian Zi	15.67	–
Jiao Gu Lan	15.70	–
Ji Li	16.42	–
Hong Jing Tian	16.72	–
Shan Yao	16.93	–
Bai Jiang Cao	17.16	–
Guang Huo Xiang	17.24	–
Che Qian Zi	17.30	–
Pi Pa Ye	17.35	–
Fu Pen Zi	17.51	–
Huang Bai	18.10	–
Yan Hu Suo	18.21	–
Ren Dong Teng	18.48	–
Tian Hua Fen	19.11	–
Lian Qiao	19.74	–
Zhu Ling	19.79	–
Chuan Xiong	19.92	–
Jin Yin Hua	20.82	–
Wu Wei Zi	21.33	–
Gu Sui Bu	22.24	–
Zhi Ke	22.82	–
Di Gu Pi	22.92	–
Ma Huang	23.30	–
Huang Lian	23.55	–
Ji Xue Teng	23.84	–
Ban Lan Gen	24.06	–
He Huan Pi	24.64	–
Bai Shao Yao	25.10	–
Suan Zao Ren	25.68	–
Gua Lou	26.04	–
Gou Teng	26.33	–
Tu Fu Ling	26.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Cao	26.68	–
Chuang Mu Xiang	26.77	–
Yuan Zhi	27.03	–
Sheng Jiang	27.89	–
Zhe Bei Mu	28.66	–
Cang Zhu	28.74	–
(Fen) Bi Xie	29.25	–
Pu Gong Ying	30.35	–
Ren Shen	30.38	–
She Gan	30.41	–
Gou Qi Zi	30.62	–
Fu Ling	30.63	–
Lu Gen	30.67	–
Rou Gui	30.84	–
Ma Huang Gen	31.28	–
Mang Xiao	32.55	–
Ze Lan	32.66	–
Zhi Gan Cao	33.18	–
Huo Ma Ren	33.46	–
Mu Zei	33.67	–
Cang Er Zi	35.02	–
Sang Ye	35.03	–
Xi Xian Cao	35.39	–
Gui Zhi	35.76	–
Yi Yi Ren	36.39	–
Tao Ren	36.91	–
Hong Hua	37.11	–
Chen Pi	37.17	–
Chi Shao (Yao)	37.84	–
Sang Zhi	38.27	–
Ban Zhi Lian	38.43	–
Xie Bai	39.17	–
Tai Zi Shen	40.42	–
Nü Zhen Zi	40.56	–
Shan Yu Rou	41.09	–
Bai Zi Ren	41.14	–
E Zhu	41.76	–
Jie Geng	42.55	–
Jiang Huang	42.60	–
Wu Mei	42.65	–
Yu Zhu	43.03	–
Zhu Ru	43.17	–
Ci Wu Jia	43.45	–
Ze Xie	43.47	–
Huang Qin	44.16	–
Dang Gui	45.03	–
San Qi	45.92	–
Ban Xia (Jiang)	46.28	–
Niu Bang Zi	47.44	–
Fu Xiao Mai	48.16	–
Dang Gui Wei	48.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Qiang Huo* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62719	62719	0.00	6.78
62720	62720	0.00	6.07

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Qin Jiao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60222-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Qin Jiao; Gentianae macrophyllae radix

Special notes

When selecting the *Qin Jiao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Qin Jiao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Qin Jiao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Qin Jiao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Qin Jiao	G113H1059921	62735	40	from supplier
PhytoComm	Qin Jiao	G113H1059921	62736	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Qin Jiao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Qin Jiao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Qin Jiao	G113H1059921	62735†	20
PhytoComm	Qin Jiao	G113H1059921	62736†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Qin Jiao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Qin Jiao	G113H1059121	1
PhytoComm	Qin Jiao	G113H1059521	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Qin Jiao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Qin Jiao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	2	855

The substance/substance group *Qin Jiao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Qin Jiao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Long Dan (Cao)	7.15	–
Zhi Gan Cao	12.25	–
Bai Zhu	12.33	–
Huang Qi	13.44	–
Jie Geng	14.33	–
Sang Zhi	18.48	–
Bai He	19.43	–
Chuan Mu Tong	19.84	–
(Huai) Niu Xi	20.89	–
Tian Hua Fen	22.28	–
Chuan Lian Zi	22.75	–
Dang Gui	23.31	–
Chuan Niu Xi	24.93	–
Chi Shao (Yao)	26.16	–
Di Gu Pi	26.24	–
Yuan Zhi	26.95	–
Bai Zhi	27.53	–
Mu Gua	27.72	–
Mang Xiao	29.23	–
Lai Fu Zi	29.36	–
Sha Shen (Bei)	29.94	–
Mi Huan Jun	29.96	–
Zi Su Zi	30.16	–
Gua Lou	30.76	–
Bing Lang	31.51	–
(Shi) Chang Pu	34.03	–
Gan Jiang	34.78	–
E Zhu	34.81	–
Lian Qiao	35.84	–
Mai Ya	35.98	–
Da Zao	36.24	–
Shan Yao	37.02	–
Ji Li	37.93	–
Chen Pi	39.14	–
Cang Er Zi	41.86	–
Ban Zhi Lian	41.99	–
Fang Feng	42.28	–
Ku Shen	42.67	–
Tu Fu Ling	44.81	–
Zhi Ke	45.57	–
Niu Bang Zi	45.76	–
Bai Shao Yao	45.95	–
Yan Hu Suo	46.18	–
Lian Zi	46.30	–
Bai Xian Pi	48.96	–
Jiang Huang	48.97	–
Gan Cao	49.03	–
Ju Hua	49.67	–
Chai Hu	49.98	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/substance group *Qin Jiao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62735	62735	0.00	8.66
62736	62736	0.00	7.15

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Qing Hao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60198-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Qing Hao; Artemisiae apiacae herba

Special notes

When selecting the *Qing Hao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Qing Hao	1	0	0

Second-stage model

For differentiation of the substance/substance group *Qing Hao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Qing Hao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Qing Hao	G031HS145TR1	63023	40	from supplier
PhytoComm	Qing Hao	G031HS145TR1	63024	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Qing Hao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Qing Hao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Qing Hao	G031HS145TR1	63023†	20
PhytoComm	Qing Hao	G031HS145TR1	63024†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Qing Hao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Qing Hao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Qing Hao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	2	40	0	12 435
Type C	0	0	0	857

The substance/substance group *Qing Hao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9857 % (> 99.9560 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Qing Hao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Qiang Huo	6.38	–
Qing Pi	7.68	–
Du Zhong	7.95	–
Zi Hua Di Ding	8.09	–
Jiao Gu Lan	8.86	–
Mao Dong Qing	9.48	–
Huang Bai	9.97	–
Dan Shen	10.58	–
Jing Jie	10.93	–
Chai Hu	10.95	–
Bo He	12.39	–
Yu Jin	12.75	–
Fu Zi	13.33	–
Ce Bai Ye	13.76	–
Yan Hu Suo	14.21	–
Hou Po	14.41	–
Sha Ren	14.64	–
Hong Jing Tian	14.81	–
Shan Yao	15.26	–
Fu Pen Zi	15.49	–
Pi Pa Ye	16.20	–
Bai Jiang Cao	16.85	–
Dan Dou Chi	16.91	–
Ling Zhi	16.98	–
Yin Yang Huo	17.41	–
Jin Yin Hua	17.96	–
Tian Hua Fen	18.52	–
Chuan Xiong	18.64	–
Wu Wei Zi	18.93	–
Huang Lian	18.97	–
Bai Xian Pi	19.12	–
Gu Sui Bu	19.36	–
Zhi Ke	19.81	–
Ye Jiao Teng	19.88	–
Sang Ye	20.00	–
Ma Huang	20.32	–
Ren Dong Teng	20.53	–
Ji Li	20.66	–
Ban Lan Gen	21.10	–
Gan Cao	21.51	–
Guang Huo Xiang	21.63	–
Che Qian Zi	21.87	–
Zhu Ling	22.82	–
Shen Qu	23.50	–
She Gan	23.64	–
Tu Fu Ling	23.69	–
Zhe Bei Mu	24.15	–
Bai Shao Yao	24.23	–
Lian Qiao	24.55	–
Ren Shen	24.87	–
Di Gu Pi	24.88	–
Lian Zi	25.20	–
Xi Xian Cao	25.43	–
Gua Lou	25.83	–
Ji Xue Teng	25.99	–
Pu Gong Ying	26.42	–
Hong Hua	26.81	–
He Huan Pi	27.65	–
Cang Zhu	28.29	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Qi Zi	28.30	–
Chuang Mu Xiang	28.92	–
Mu Zei	28.94	–
Zhi Gan Cao	29.09	–
Ze Lan	30.13	–
Yuan Zhi	30.63	–
Sheng Jiang	31.14	–
Lu Gen	31.14	–
Cang Er Zi	31.50	–
Suan Zao Ren	31.71	–
Wu Mei	31.84	–
(Fen) Bi Xie	31.87	–
Mang Xiao	32.04	–
Nü Zhen Zi	33.18	–
Gou Teng	33.20	–
Fu Ling	34.33	–
Shan Yu Rou	34.78	–
Ze Xie	35.57	–
Xie Bai	35.85	–
Jie Geng	36.02	–
Huo Ma Ren	36.42	–
Chen Pi	37.84	–
Tao Ren	38.35	–
Chi Shao (Yao)	38.75	–
Huang Qin	39.45	–
Zhi Shi	39.54	–
Bai Zi Ren	40.25	–
Ku Shen	41.27	–
Dang Gui	41.32	–
Tai Zi Shen	41.51	–
Ma Huang Gen	41.67	–
Jiang Huang	41.76	–
Gui Zhi	42.02	–
Sang Zhi	43.55	–
San Qi	43.65	–
Rou Gui	43.83	–
Xin Yi	44.52	–
E Zhu	44.79	–
Yu Zhu	45.00	–
Ban Zhi Lian	45.37	–
Yi Mu Cao	46.34	–
Yi Yi Ren	46.61	–
Zhu Ru	47.23	–
Yin Chen Hao	48.24	–
Ban Xia (Jiang)	49.03	–
Ci Wu Jia	49.19	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Qing Hao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
63023	63023	0.00	6.38
63024	63024	0.00	7.04

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Qing Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50371-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Qing Pi; Citri reticulatae viride pericarpium

Special notes

When selecting the *Qing Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Qing Pi	1	0	2

Second-stage model

For differentiation of the substance/substance group *Qing Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Qing Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Qing Pi	G074HS361SK1	62671	40	from supplier
PhytoComm	Qing Pi	G074HS361SK1	62672	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Qing Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Qing Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Qing Pi	G074HS361SK1	62671†	20
PhytoComm	Qing Pi	G074HS361SK1	62672†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Qing Pi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Qing Pi	G074H0809621	1
Phytocomm	Qing Pi	g074h100713	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Qing Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Qing Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	2	855

The substance/substance group *Qing Pi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Qing Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Qing Hao	5.46	–
Huang Bai	9.88	–
Zi Hua Di Ding	12.04	–
Zhi Ke	12.65	–
Gan Cao	13.91	–
Sang Ye	14.39	–
Shan Yao	14.72	–
Jiao Gu Lan	15.70	–
Wu Wei Zi	15.83	–
Du Zhong	16.57	–
Qiang Huo	16.95	–
Mao Dong Qing	17.84	–
Fu Zi	18.16	–
Dan Dou Chi	18.63	–
Chai Hu	18.64	–
Jin Yin Hua	18.78	–
Sha Ren	20.00	–
Dan Shen	20.20	–
Hou Po	20.77	–
Huang Lian	20.86	–
Fu Pen Zi	21.16	–
Zhe Bei Mu	21.31	–
Jing Jie	21.45	–
Hong Hua	21.52	–
Bo He	22.47	–
Ma Huang	22.51	–
Hong Jing Tian	22.95	–
Yan Hu Suo	23.50	–
Ban Lan Gen	23.67	–
Ce Bai Ye	23.90	–
Wu Mei	24.00	–
Yin Yang Huo	24.41	–
Ye Jiao Teng	24.95	–
Zhi Gan Cao	25.14	–
Bai Jiang Cao	25.18	–
She Gan	25.32	–
Suan Zao Ren	25.77	–
Chuan Xiong	26.07	–
Lian Qiao	26.34	–
Guang Huo Xiang	26.49	–
Chuang Mu Xiang	26.59	–
Pi Pa Ye	26.71	–
Cang Zhu	27.78	–
Ren Dong Teng	28.60	–
Shan Yu Rou	29.03	–
Gou Qi Zi	29.19	–
Tian Hua Fen	29.42	–
Cang Er Zi	29.64	–
Gu Sui Bu	30.09	–
Xi Xian Cao	30.11	–
Ji Li	30.20	–
Xie Bai	30.31	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Shen	30.70	–
Ku Shen	31.02	–
Bai Shao Yao	31.64	–
He Huan Pi	31.64	–
Che Qian Zi	31.74	–
Mang Xiao	31.79	–
Tu Fu Ling	31.95	–
Gua Lou	32.43	–
Yu Jin	32.49	–
Mu Zei	32.62	–
Di Gu Pi	32.99	–
Jie Geng	33.24	–
Ling Zhi	33.26	–
Lu Gen	33.62	–
Bai Xian Pi	33.69	–
Nü Zhen Zi	35.14	–
Ze Lan	35.32	–
Ze Xie	35.54	–
Lian Zi	35.87	–
Pu Gong Ying	35.88	–
Zhu Ling	36.14	–
Dang Gui	37.76	–
Sheng Jiang	38.16	–
Huang Qin	38.22	–
Chen Pi	40.36	–
San Qi	40.51	–
Tao Ren	41.07	–
Tai Zi Shen	41.95	–
Huo Ma Ren	42.15	–
Ji Xue Teng	42.16	–
Shen Qu	43.05	–
(Fen) Bi Xie	43.22	–
Yuan Zhi	43.38	–
Bai Zi Ren	43.56	–
Zhi Shi	43.72	–
Chi Shao (Yao)	43.91	–
Xin Yi	44.10	–
Yu Zhu	44.46	–
Gou Teng	45.63	–
Jiang Huang	45.99	–
Yin Chen Hao	46.98	–
Yi Mu Cao	47.97	–
Fu Ling	48.02	–
Zhu Ru	49.65	–
Gui Zhi	50.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Qing Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62671	62671	0.00	6.05
62672	62672	0.00	5.46

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ren Dong Teng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60104-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ren Dong Teng; Lonicerae japonicae caulis

Special notes

When selecting the *Ren Dong Teng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ren Dong Teng	1	0	0

Second-stage model

For differentiation of the substance/substance group *Ren Dong Teng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ren Dong Teng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ren Dong Teng	G297HS144SK1	62695	40	from supplier
PhytoComm	Ren Dong Teng	G297HS144SK1	62696	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ren Dong Teng*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ren Dong Teng*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ren Dong Teng	G297HS144SK1	62695†	20
PhytoComm	Ren Dong Teng	G297HS144SK1	62696†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Ren Dong Teng*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ren Dong Teng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ren Dong Teng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	79	1	24 517
Type B	3	39	1	12 434
Type C	0	0	0	857

The substance/substance group *Ren Dong Teng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9786 % (> 99.9636 %)	98.7500 % (> 95.0000 %)
Type B	99.9571 % (> 99.9274 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ren Dong Teng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Ling Zhi	5.32	–
Ce Bai Ye	5.80	–
Shen Qu	5.82	–
(Fen) Bi Xie	5.96	–
Bai Xian Pi	6.76	–
She Gan	6.89	–
Gu Sui Bu	8.31	–
Zhu Ling	8.44	–
Lu Gen	8.52	–
Ye Jiao Teng	8.59	–
Dan Dou Chi	8.64	–
Lian Zi	9.19	–
Ji Xue Teng	9.26	–
Yu Jin	9.43	–
Ji Li	9.62	–
Guang Huo Xiang	9.65	–
Chai Hu	9.97	–
Rou Gui	9.98	–
Huo Ma Ren	11.06	–
Chuan Xiong	11.23	–
Hou Po	11.30	–
Gou Teng	11.35	–
Che Qian Zi	11.42	–
Shan Yao	11.63	–
Ma Huang Gen	11.91	–
Lian Qiao	11.94	–
Mao Dong Qing	12.06	–
Pi Pa Ye	12.14	–
He Huan Pi	12.24	–
Suan Zao Ren	12.36	–
Bai Shao Yao	12.48	–
Fu Pen Zi	12.55	–
Tu Fu Ling	12.72	–
Fu Ling	13.24	–
Dan Shen	13.55	–
Yan Hu Suo	13.62	–
Yin Yang Huo	13.77	–
Mu Zei	13.84	–
Tian Hua Fen	14.10	–
Fu Zi	14.40	–
Sheng Jiang	14.59	–
Gui Zhi	15.68	–
Yi Yi Ren	17.06	–
Gua Lou	17.49	–
Tai Zi Shen	17.66	–
Zhe Bei Mu	17.81	–
Qiang Huo	18.03	–
Jiao Gu Lan	18.49	–
Tao Ren	18.53	–
Fo Shou	18.56	–
Gan Cao	18.73	–
Bai Zi Ren	19.06	–
Jin Yin Hua	19.06	–
Hong Jing Tian	19.59	–
Ze Xie	19.83	–
Ma Huang	20.03	–
Ban Lan Gen	20.65	–
Jing Jie	20.93	–
Ci Wu Jia	21.57	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Di Gu Pi	21.85	–
Chen Pi	22.65	–
Ban Xia (Jiang)	23.24	–
Zi Hua Di Ding	23.34	–
Fu Xiao Mai	23.36	–
Yuan Zhi	23.68	–
Cang Zhu	24.01	–
Zhu Ru	24.54	–
Zhi Gan Cao	25.42	–
Fu Shen	25.83	–
Qing Pi	26.36	–
Yu Zhu	26.63	–
Sha Ren	26.67	–
Du Zhong	26.93	–
Zhi Ke	27.00	–
Ren Shen	27.42	–
Wu Wei Zi	27.61	–
Lai Fu Zi	27.88	–
Bo He	28.89	–
Dang Gui Wei	30.21	–
Huang Lian	30.58	–
Huang Bai	31.29	–
Chuang Mu Xiang	31.48	–
San Qi	31.58	–
Mang Xiao	31.96	–
Gou Qi Zi	32.23	–
Dang Gui	32.48	–
Jie Geng	32.55	–
Huang Qin	36.41	–
Xie Bai	36.51	–
Sang Zhi	37.67	–
Ku Shen	38.59	–
Shan Yu Rou	39.05	–
Qing Hao	40.20	–
Chi Shao (Yao)	41.32	–
Long Yan Rou	42.44	–
Bai Jiang Cao	42.61	–
Zhi Mu	43.45	–
Cang Er Zi	47.08	–
Ban Zhi Lian	47.36	–
Mu Gua	47.41	–
Hong Hua	47.88	–
Bai Zhu	50.02	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ren Dong Teng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62695	62695	0.00	5.96
62696	62696	0.00	5.32

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ren Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60984-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ren Shen; Ginseng radix et rhizoma

Special notes

When selecting the *Ren Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ren Shen	2	0	1

Second-stage model

For differentiation of the substance/substance group *Ren Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ren Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ren Shen	G115HS003TG1	62881	40	from supplier
PhytoComm	Ren Shen	G115HS003TG1	62882	40	from supplier
PhytoComm	Ren Shen	G115H0562922	63001	40	from supplier
PhytoComm	Ren Shen	G115H0562922	63002	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ren Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ren Shen*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ren Shen	G115HS003TG1	62881†	20
PhytoComm	Ren Shen	G115HS003TG1	62882†	20
PhytoComm	Ren Shen	G115H0562922	63001†	20
PhytoComm	Ren Shen	G115H0562922	63002†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Ren Shen*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ren Shen	g115ho562121	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ren Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ren Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	1	856

The substance/substance group *Ren Shen* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ren Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Long Yan Rou	13.66	–
Ye Jiao Teng	16.00	–
Gou Teng	16.50	–
Yu Zhu	16.77	–
Shan Yu Rou	16.80	–
Bai Shao Yao	17.05	–
Chen Pi	17.09	–
Gui Zhi	17.36	–
Zhe Bei Mu	17.87	–
Jin Yin Hua	18.03	–
Ma Huang	18.11	–
Dang Gui	19.21	–
Gua Lou	19.30	–
Lian Qiao	19.31	–
Fo Shou	19.45	–
Tian Hua Fen	19.66	–
Jiao Gu Lan	20.10	–
Rou Gui	20.26	–
Yuan Zhi	20.49	–
Cang Zhu	21.06	–
Di Gu Pi	21.08	–
Lu Gen	21.09	–
Tai Zi Shen	21.28	–
Jie Geng	21.37	–
Zhi Ke	21.65	–
Tao Ren	21.96	–
Ban Xia (Jiang)	22.08	–
Ji Xue Teng	22.48	–
San Qi	22.74	–
He Huan Pi	22.95	–
Mu Zei	23.71	–
Pi Pa Ye	23.80	–
Huang Qin	24.01	–
Ze Xie	24.06	–
Fu Shen	24.61	–
Ji Li	24.69	–
Dang Gui Wei	24.93	–
Ren Dong Teng	25.10	–
Shen Qu	25.18	–
Zhi Mu	25.33	–
Bai Xian Pi	25.65	–
Zhu Ru	26.16	–
Shan Yao	26.26	–
Gu Sui Bu	26.26	–
Tu Fu Ling	26.48	–
Yin Yang Huo	26.58	–
Gou Qi Zi	26.81	–
Ling Zhi	26.93	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yan Hu Suo	27.10	–
Ma Huang Gen	27.10	–
Bai Zi Ren	27.12	–
Ci Wu Jia	27.17	–
Hou Po	27.33	–
Wu Wei Zi	27.51	–
Dan Shen	27.90	–
Ku Shen	27.98	–
Huo Ma Ren	28.15	–
Ban Lan Gen	28.56	–
Fu Zi	28.84	–
Fu Ling	28.88	–
Mang Xiao	29.06	–
Fu Xiao Mai	29.13	–
Hong Jing Tian	29.15	–
Lian Zi	29.28	–
Sheng Jiang	29.31	–
Lai Fu Zi	29.80	–
Chuang Mu Xiang	29.86	–
Gan Cao	29.98	–
Fu Pen Zi	30.47	–
Chuan Xiong	30.61	–
(Fen) Bi Xie	31.21	–
Mu Gua	31.38	–
Yi Yi Ren	32.70	–
Che Qian Zi	32.80	–
Dan Dou Chi	32.86	–
Sang Zhi	33.43	–
Guang Huo Xiang	34.32	–
Suan Zao Ren	34.59	–
She Gan	34.86	–
Yu Jin	35.00	–
Chi Shao (Yao)	35.46	–
Chai Hu	35.59	–
Hong Hua	36.15	–
Mai Men Dong	36.76	–
Xie Bai	37.24	–
Huang Qi	37.58	–
Huang Lian	37.76	–
Qing Pi	38.73	–
Chuan Lian Zi	40.07	–
Ce Bai Ye	40.39	–
Jing Jie	40.90	–
Qiang Huo	40.95	–
Zi Hua Di Ding	41.72	–
Zhu Ling	42.62	–
Mao Dong Qing	42.88	–
Zhi Gan Cao	45.55	–
Sha Ren	45.63	–
Bo He	47.04	–
Sang Ye	47.67	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ren Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62881	62881	0.00	13.66
62882	62882	0.00	15.02
63001	63001	0.00	16.95
63002	63002	0.00	16.80

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Rou Cong Rong**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60182-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Rou Cong Rong; Cistanchis herba

Special notes

When selecting the *Rou Cong Rong* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Rou Cong Rong	1	0	0

Second-stage model

For differentiation of the substance/substance group *Rou Cong Rong* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Rou Cong Rong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Rou Cong Rong	G072HS108SG1	62667	40	from supplier
PhytoComm	Rou Cong Rong	G072HS108SG1	62668	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Rou Cong Rong*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Rou Cong Rong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Rou Cong Rong	G072HS108SG1	62667†	20
PhytoComm	Rou Cong Rong	G072HS108SG1	62668†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Rou Cong Rong*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Rou Cong Rong* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Rou Cong Rong* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	0	857

The substance/substance group *Rou Cong Rong* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Rou Cong Rong* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Zhi Shi	15.93	–
Bai Hua She She Cao	16.02	–
Hu Zhang	17.67	–
Han Lian Cao	19.17	–
Xin Yi	20.25	–
Yu Xing Cao	21.93	–
Huang Lian	23.87	–
Sha Ren	23.96	–
Pu Gong Ying	24.01	–
Mang Xiao	25.78	–
Sang Ye	25.82	–
Yi Mu Cao	26.54	–
Xi Xian Cao	27.46	–
Gu Sui Bu	28.89	–
Xia Ku Cao	29.33	–
Gou Teng	29.41	–
Ze Lan	31.31	–
Tu Fu Ling	31.54	–
Nü Zhen Zi	32.44	–
Xiang Fu	32.59	–
Shan Yao	34.87	–
Sang Ji Shend	35.64	–
Yin Chen Hao	37.02	–
Ju Hua	37.18	–
Xiao Hui Xiang	39.07	–
Jiang Huang	39.08	–
Wu Zhu Yu	39.95	–
He Huan Pi	41.43	–
Cang Er Zi	41.43	–
Dan Zhu Ye	42.26	–
Jing Jie	46.17	–
Gan Cao	46.72	–
Wu Mei	48.25	–
Wu Yao	48.29	–
Qiang Huo	48.68	–
Sang Bai Pi	49.99	–
Wu Wei Zi	50.83	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Rou Cong Rong* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62667	62667	0.00	15.93
62668	62668	0.00	16.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Rou Gui**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60169-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Rou Gui; Cinnamomi cassiae cortex

Special notes

When selecting the *Rou Gui* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Rou Gui	3	0	3

Second-stage model

For differentiation of the substance/substance group *Rou Gui* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Rou Gui*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Rou Gui	G069HS204RN1	62271	40	from supplier
PhytoComm	Rou Gui	G069HS204RN1	62272	40	from supplier
PhytoComm	Rou Gui	G069HS204RM1	62443	40	from supplier
PhytoComm	Rou Gui	G069HS204RM1	62444	40	from supplier
PhytoComm	Rou Gui	G069HS204TH1	62917	40	from supplier
PhytoComm	Rou Gui	G069HS204TH1	62918	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Rou Gui*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Rou Gui*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Rou Gui	G069HS204RN1	62271†	20
PhytoComm	Rou Gui	G069HS204RN1	62272†	20
PhytoComm	Rou Gui	G069HS204RM1	62443†	20
PhytoComm	Rou Gui	G069HS204RM1	62444†	20
PhytoComm	Rou Gui	G069HS204TH1	62917†	20
PhytoComm	Rou Gui	G069HS204TH1	62918†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 5 *Apo-Ident* customers from 4 batches from the substance/substance group *Rou Gui*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Rou Gui	g069h0617123	1
PhytoComm	Rou Gui	G069H0617123	2
Phytocomm	Rou Gui	G069H0617323	1
PhytoComm	Rou Gui	G069H0617323	1
Phytocomm	Rou Gui	H0617022	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Rou Gui* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Rou Gui* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	9	239	1	24 351
Type B	24	115	5	12 333
Type C	1	0	6	850

The substance/substance group *Rou Gui* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9726 % (> 99.9576 %)	99.5833 % (> 98.3333 %)
Type B	99.8476 % (> 99.8177 %)	95.8333 % (> 93.3333 %)
Type C	99.7674 % (> 99.1804 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Rou Gui* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Gui Zhi	6.77	–
Ma Huang Gen	6.87	–
Sheng Jiang	7.66	–
Yi Yi Ren	8.15	–
He Huan Pi	8.20	–
Fu Ling	8.47	–
Ban Xia (Jiang)	9.09	–
Gou Teng	9.34	–
Fu Shen	10.40	–
Mu Zei	10.41	–
Ji Xue Teng	10.57	–
Di Gu Pi	11.60	–
Fu Xiao Mai	11.67	–
Fo Shou	13.16	–
Tao Ren	13.62	–
Zhu Ru	15.34	–
Ye Jiao Teng	15.49	–
Bai Xian Pi	15.53	–
Lai Fu Zi	15.58	–
Ci Wu Jia	15.75	–
Bai Shao Yao	16.19	–
Tai Zi Shen	16.55	–
Huo Ma Ren	16.68	–
Lian Zi	17.34	–
Lu Gen	17.82	–
Bai Zi Ren	18.30	–
Gu Sui Bu	18.44	–
Ling Zhi	19.05	–
Shen Qu	19.58	–
Ji Li	20.06	–
Zhe Bei Mu	21.81	–
Ren Dong Teng	23.55	–
Ze Xie	23.74	–
Tu Fu Ling	24.58	–
Gua Lou	24.94	–
Tian Hua Fen	25.33	–
Suan Zao Ren	25.55	–
Zhu Ling	26.06	–
(Fen) Bi Xie	26.98	–
Chuan Xiong	27.49	–
Chen Pi	27.75	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
She Gan	28.09	–
Pi Pa Ye	28.75	–
Yuan Zhi	28.84	–
Dan Dou Chi	29.32	–
Yu Jin	29.47	–
Lian Qiao	29.81	–
Yan Hu Suo	31.50	–
Dan Shen	31.79	–
Mang Xiao	32.24	–
Jin Yin Hua	33.48	–
Shan Yao	33.88	–
Zhi Mu	33.95	–
Jiao Gu Lan	34.74	–
Ren Shen	34.86	–
Yu Zhu	34.96	–
Zhi Ke	35.25	–
Jie Geng	35.43	–
Dang Gui	35.76	–
Cang Zhu	35.85	–
Ce Bai Ye	36.07	–
Yin Yang Huo	36.19	–
Hong Jing Tian	37.46	–
Shan Yu Rou	37.86	–
Ban Lan Gen	37.98	–
Huang Qin	37.99	–
Chai Hu	38.68	–
Fu Pen Zi	39.04	–
Ma Huang	39.07	–
Mao Dong Qing	39.27	–
Fu Zi	39.57	–
Guang Huo Xiang	40.07	–
Dang Gui Wei	41.08	–
San Qi	42.19	–
Hou Po	45.12	–
Gan Cao	45.82	–
Mu Gua	46.11	–
Qiang Huo	46.64	–
Sang Zhi	48.09	–
Wu Wei Zi	48.75	–
Gou Qi Zi	48.76	–
Long Yan Rou	49.08	–
Jing Jie	49.13	–
Huang Lian	49.75	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Rou Gui* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62271	62271	0.00	7.04
62272	62272	0.00	6.86
62443	62443	0.00	6.77
62444	62444	0.00	6.87
62917	62917	0.00	8.55
62918	62918	0.00	7.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**San Qi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60415-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

San Qi; Notoginseng radix

Special notes

When selecting the *San Qi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
San Qi	1	0	1

Second-stage model

For differentiation of the substance/substance group *San Qi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *San Qi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	San Qi	G203HS030SL1	62701	40	from supplier
PhytoComm	San Qi	G203HS030SL1	62702	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *San Qi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *San Qi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	San Qi	G203HS030SL1	62701†	20
PhytoComm	San Qi	G203HS030SL1	62702†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *San Qi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	San Qi	g203h0312021	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *San Qi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *San Qi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	1	0	1	855

The substance/substance group *San Qi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8450 % (> 99.2633 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *San Qi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ku Shen	5.13	–
Shan Yao	9.07	–
Lian Qiao	9.23	–
Pi Pa Ye	9.50	–
Yuan Zhi	9.58	–
Bai Shao Yao	10.61	–
Zhi Mu	10.73	–
Ji Li	11.56	–
Cang Zhu	11.84	–
Bai Zi Ren	11.86	–
Zhu Ru	14.19	–
Lai Fu Zi	14.25	–
Suan Zao Ren	14.33	–
Zhe Bei Mu	14.61	–
Ren Shen	14.75	–
Jiao Gu Lan	15.49	–
Shan Yu Rou	15.67	–
Huo Ma Ren	15.81	–
Jie Geng	15.84	–
Ling Zhi	16.07	–
Tian Hua Fen	16.23	–
Gua Lou	16.65	–
Xie Bai	16.76	–
Lian Zi	16.81	–
Jin Yin Hua	16.86	–
Dang Gui	17.13	–
Mu Gua	17.37	–
Dang Gui Wei	17.61	–
Shen Qu	18.85	–
Sheng Jiang	19.11	–
Gou Qi Zi	19.13	–
Di Gu Pi	19.22	–
Mu Zei	19.37	–
Ren Dong Teng	19.53	–
Fu Zi	19.77	–
Yu Zhu	20.11	–
Chen Pi	20.39	–
Wu Wei Zi	20.41	–
Hou Po	20.60	–
Fu Ling	20.64	–
Tai Zi Shen	20.67	–
Mao Dong Qing	20.94	–
He Huan Pi	21.22	–
Chuan Xiong	21.25	–
Ye Jiao Teng	21.80	–
Chuang Mu Xiang	22.29	–
Ze Xie	22.77	–
Ban Lan Gen	24.83	–
Gui Zhi	24.95	–
Ci Wu Jia	25.06	–
Dan Dou Chi	25.07	–
Zhi Ke	25.37	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Shen	26.42	–
Ji Xue Teng	26.76	–
Fu Xiao Mai	26.83	–
Yi Yi Ren	27.34	–
Tu Fu Ling	27.41	–
Lu Gen	28.02	–
Che Qian Zi	28.20	–
Gou Teng	28.43	–
Gu Sui Bu	28.63	–
Fu Pen Zi	29.04	–
Yan Hu Suo	29.37	–
Tao Ren	29.42	–
Chai Hu	30.20	–
Yu Jin	30.30	–
Ban Xia (Jiang)	30.53	–
Rou Gui	30.76	–
Fo Shou	30.85	–
Guang Huo Xiang	30.85	–
Bai Xian Pi	31.59	–
Huang Qin	32.20	–
Gan Cao	32.42	–
Mang Xiao	33.79	–
Long Yan Rou	34.05	–
Ce Bai Ye	34.67	–
Yin Yang Huo	35.31	–
Zhi Gan Cao	35.49	–
Zhu Ling	35.52	–
Hong Jing Tian	35.67	–
Qing Pi	36.07	–
Qiang Huo	36.46	–
She Gan	38.34	–
Ma Huang	38.93	–
Cang Er Zi	40.27	–
(Fen) Bi Xie	40.56	–
Sang Zhi	41.11	–
Ma Huang Gen	41.44	–
Mai Men Dong	41.44	–
Chi Shao (Yao)	41.88	–
Du Zhong	43.07	–
Zi Hua Di Ding	43.25	–
Huang Lian	43.66	–
Bo He	44.41	–
Hong Hua	46.62	–
Sha Ren	48.50	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *San Qi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62701	62701	0.00	5.13
62702	62702	0.00	5.24

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sang Bai Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50287-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sang Bai Pi; Mori cortex

Special notes

When selecting the *Sang Bai Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sang Bai Pi	1	0	2

Second-stage model

For differentiation of the substance/substance group *Sang Bai Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sang Bai Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sang Bai Pi	G161H1042921	62883	40	from supplier
PhytoComm	Sang Bai Pi	G161H1042921	62884	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Sang Bai Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Sang Bai Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sang Bai Pi	G161H1042921	62883†	20
PhytoComm	Sang Bai Pi	G161H1042921	62884†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Sang Bai Pi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Sang Bai Pi	G161H1042321	1
PhytoComm	Sang Bai Pi	G161H1042321	2
PhytoComm	Sang Bai Pi	G161H1042322	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sang Bai Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sang Bai Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	1	0	4	852

The substance/substance group *Sang Bai Pi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.7674 % (> 99.1809 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sang Bai Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yin Chen Hao	10.89	–
Gu Sui Bu	11.81	–
Dan Zhu Ye	12.25	–
Xuan Fu Hua	15.67	–
He Huan Pi	15.78	–
Sang Ji Shend	18.01	–
Bu Gu Zhi	18.77	–
He Shou Wu	19.78	–
Chen Pi	20.12	–
Wu Zhu Yu	20.76	–
Sha Ren	24.14	–
Tu Si Zi	24.72	–
Mang Xiao	26.58	–
Jiang Huang	27.11	–
Nü Zhen Zi	27.89	–
Jing Jie	27.96	–
(Shi) Chang Pu	29.06	–
(Bai) Dou Kou	33.58	–
Wu Yao	33.87	–
Yi Mu Cao	34.11	–
Ge Gen	35.14	–
Ji Li	35.66	–
Jin Qian Cao	35.84	–
Du Zhong	36.24	–
Hong Jing Tian	38.23	–
Shan Yao	40.86	–
Mu Dan Pi	41.22	–
Xiao Hui Xiang	41.37	–
Gou Teng	41.41	–
Hu Zhang	42.29	–
Sang Ye	44.80	–
Shan Yu Rou	45.31	–
Fang Feng	45.76	–
Ju Hua	45.93	–
Wu Jia Pi	46.56	–
Ku Shen	48.20	–
E Zhu	48.22	–
Zhi Shi	49.41	–
Hua Shi	50.48	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sang Bai Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62883	62883	0.00	11.29
62884	62884	0.00	10.89

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sang Ji Shend**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60142-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sang Ji Shend; Taxilli herba

Special notes

When selecting the *Sang Ji Shend* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sang Ji Shend	3	0	3

Second-stage model

For differentiation of the substance/substance group *Sang Ji Shend* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sang Ji Shend*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sang Ji Shend	G146H1045821	62345	40	from supplier
PhytoComm	Sang Ji Shend	G146H1045821	62346	40	from supplier
PhytoComm	Sang Ji Shend	G146H1045822	62543	40	from supplier
PhytoComm	Sang Ji Shend	G146H1045822	62544	40	from supplier
PhytoComm	Sang Ji Shend	G146H1045922	62865	40	from supplier
PhytoComm	Sang Ji Shend	G146H1045922	62866	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Sang Ji Shend*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Sang Ji Shend*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sang Ji Shend	G146H1045821	62345†	20
PhytoComm	Sang Ji Shend	G146H1045821	62346†	20
PhytoComm	Sang Ji Shend	G146H1045822	62543†	20
PhytoComm	Sang Ji Shend	G146H1045822	62544†	20
PhytoComm	Sang Ji Shend	G146H1045922	62865†	20
PhytoComm	Sang Ji Shend	G146H1045922	62866†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 5 *Apo-Ident* customers from 3 batches from the substance/substance group *Sang Ji Shend*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Sang Ji Shend	g146h1045222	1
PhytoComm	Sang Ji Shend	G146H1045322	1
Phytocomm	Sang Ji Shend	G146H1045522	2
PhytoComm	Sang Ji Shend	G146H1045522	2

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sang Ji Shend* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sang Ji Shend* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	2	4	851

The substance/substance group *Sang Ji Shend* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sang Ji Shend* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
He Shou Wu	10.89	–
Gu Sui Bu	11.43	–
Xuan Fu Hua	14.06	–
Yin Chen Hao	17.57	–
Sang Bai Pi	22.06	–
Dan Zhu Ye	22.54	–
Mang Xiao	23.25	–
Ge Gen	23.62	–
He Huan Pi	23.83	–
Hong Jing Tian	25.86	–
Jing Jie	27.16	–
Wu Zhu Yu	29.66	–
Bu Gu Zhi	31.92	–
Chen Pi	32.28	–
Nü Zhen Zi	33.53	–
Sha Ren	34.34	–
Gou Teng	36.65	–
Jin Qian Cao	38.87	–
Yi Mu Cao	38.97	–
Du Zhong	39.11	–
Hu Zhang	41.19	–
Jiang Huang	43.58	–
Tu Si Zi	44.03	–
Wu Yao	44.81	–
Zhi Shi	44.86	–
Sang Ye	45.78	–
Bai Hua She She Cao	46.59	–
Mu Dan Pi	46.73	–
(Shi) Chang Pu	47.53	–
Shan Yu Rou	47.72	–
(Bai) Dou Kou	48.36	–
Xia Ku Cao	48.78	–
Fang Feng	49.22	–
Xiao Hui Xiang	49.68	–
Wu Jia Pi	49.92	–
Hua Shi	49.93	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sang Ji Shend* is separated from critical neighbours in a second-stage model,

all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62345	62345	0.00	16.90
62346	62346	0.00	16.96
62543	62543	0.00	19.44
62544	62544	0.00	19.95
62865	62865	0.00	11.53
62866	62866	0.00	10.89

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sang Ye**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60079-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sang Ye; Mori albi folium

Special notes

When selecting the *Sang Ye* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sang Ye	2	0	2

Second-stage model

For differentiation of the substance/substance group *Sang Ye* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sang Ye*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sang Ye	G162HS206RQ1	62461	40	from supplier
PhytoComm	Sang Ye	G162HS206RQ1	62462	40	from supplier
PhytoComm	Sang Ye	G162H1043021	62993	40	from supplier
PhytoComm	Sang Ye	G162H1043021	62994	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Sang Ye*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Sang Ye*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sang Ye	G162HS206RQ1	62461†	20
PhytoComm	Sang Ye	G162HS206RQ1	62462†	20
PhytoComm	Sang Ye	G162H1043021	62993†	20
PhytoComm	Sang Ye	G162H1043021	62994†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Sang Ye*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Sang Ye	G162H1043121	1
Phytocomm	Sang Ye	G162H1043421	4

- 852 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sang Ye* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sang Ye* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	1	80	0	12 396
Type C	0	0	5	852

The substance/substance group *Sang Ye* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9952 % (> 99.9654 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sang Ye* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zhi Shi	9.23	–
Yin Chen Hao	9.25	–
Ze Lan	12.24	–
Bai Jiang Cao	12.26	–
Shan Yao	12.85	–
Pu Gong Ying	13.08	–
Huang Lian	13.54	–
Sha Ren	14.92	–
Nü Zhen Zi	15.48	–
Tu Fu Ling	17.15	–
Wu Mei	17.80	–
Wu Jia Pi	18.06	–
Xin Yi	18.30	–
Xi Xian Cao	18.57	–
Jing Jie	18.58	–
Yu Xing Cao	18.97	–
Qiang Huo	20.05	–
Qing Pi	21.20	–
Du Zhong	21.94	–
Huang Bai	22.12	–
Zi Hua Di Ding	22.37	–
Dan Shen	22.65	–
Mang Xiao	23.02	–
Bo He	23.30	–
Fu Zi	23.50	–
Mao Dong Qing	24.49	–
Qing Hao	24.66	–
Xian Mao	24.82	–
Zhi Ke	24.87	–
Hong Jing Tian	25.26	–
Jin Qian Cao	25.33	–
(Sheng) Di Huang	26.46	–
Hou Po	28.06	–
Wu Wei Zi	28.20	–
Chai Hu	28.23	–
Ding Xiang	28.71	–
Hong Hua	28.88	–
Jiang Huang	29.21	–
Dan Dou Chi	29.91	–
Ma Huang	29.97	–
Dan Zhu Ye	30.71	–
Yin Yang Huo	30.72	–
Yu Jin	30.85	–
Han Lian Cao	31.47	–
Hu Zhang	31.50	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	31.63	–
Cang Er Zi	32.10	–
Fu Pen Zi	32.26	–
Che Qian Zi	32.74	–
Ce Bai Ye	33.44	–
Sang Ji Shend	33.99	–
Ye Jiao Teng	34.28	–
Yi Mu Cao	34.66	–
Jin Yin Hua	35.71	–
Chuang Mu Xiang	35.76	–
Gan Cao	38.11	–
Zi Su Zi	38.40	–
Bai Hua She She Cao	38.67	–
Wang Bu Liu Xing	39.89	–
Gou Teng	40.10	–
Xiang Fu	40.29	–
Niu Bang Zi	41.01	–
Chuan Xiong	41.10	–
Suan Zao Ren	41.20	–
Shan Yu Rou	41.20	–
He Huan Pi	41.33	–
Di Gu Pi	41.39	–
Ge Gen	41.55	–
Lian Qiao	42.38	–
Bai Xian Pi	42.58	–
Zhe Bei Mu	42.77	–
Yan Hu Suo	43.77	–
Ban Lan Gen	43.93	–
Pi Pa Ye	44.72	–
Gua Lou	45.05	–
Jie Geng	45.54	–
Bu Gu Zhi	45.70	–
Shu Di (Huang)	46.40	–
Zhi Gan Cao	46.44	–
Ren Dong Teng	46.70	–
Bai Shao Yao	46.81	–
Gou Qi Zi	47.04	–
Chi Shao (Yao)	47.28	–
Huang Qin	47.67	–
Hua Shi	48.06	–
Lian Zi	48.06	–
E Zhu	48.53	–
(Shi) Chang Pu	49.03	–
Cang Zhu	49.65	–
Ling Zhi	49.78	–
Xia Ku Cao	49.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sang Ye* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62461	62461	0.00	9.23
62462	62462	0.00	9.25
62993	62993	0.00	18.08
62994	62994	0.00	18.06

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sang Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60091-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sang Zhi; Mori albae ramulus

Special notes

When selecting the *Sang Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sang Zhi	2	0	0

Second-stage model

For differentiation of the substance/substance group *Sang Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sang Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sang Zhi	G164H1044821	62473	40	from supplier
PhytoComm	Sang Zhi	G164H1044821	62474	40	from supplier
PhytoComm	Sang Zhi	G164H1044921	62895	40	from supplier
PhytoComm	Sang Zhi	G164H1044921	62896	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Sang Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Sang Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sang Zhi	G164H1044821	62473†	20
PhytoComm	Sang Zhi	G164H1044821	62474†	20
PhytoComm	Sang Zhi	G164H1044921	62895†	20
PhytoComm	Sang Zhi	G164H1044921	62896†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Sang Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sang Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sang Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	5	0	0	852

The substance/substance group *Sang Zhi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.3981 % (> 98.8154 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sang Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Chuan Lian Zi	7.57	–
Tian Hua Fen	9.46	–
Mi Huan Jun	10.99	–
Jie Geng	12.35	–
Di Gu Pi	13.12	–
Gan Jiang	13.37	–
E Zhu	13.80	–
Mai Ya	13.87	–
Bai He	14.08	–
Lai Fu Zi	15.41	–
Sha Shen (Bei)	15.74	–
Chuan Mu Tong	15.85	–
Ji Li	18.00	–
Bai Zhu	18.29	–
Huang Qi	19.17	–
Zhi Gan Cao	20.87	–
Dang Gui	20.90	–
Qin Jiao	20.96	–
Zi Su Zi	21.52	–
(Shi) Chang Pu	23.00	–
Long Dan (Cao)	23.15	–
Cang Er Zi	23.61	–
Yan Hu Suo	23.98	–
Mang Xiao	24.79	–
Ban Zhi Lian	25.17	–
Gua Lou	25.19	–
Shan Yao	26.12	–
(Huai) Niu Xi	26.71	–
Chi Shao (Yao)	28.13	–
Jiang Huang	28.50	–
Mu Gua	29.72	–
Bai Zhi	30.09	–
Da Zao	30.67	–
Lian Zi	32.60	–
Bai Xian Pi	33.79	–
Lian Qiao	33.95	–
Chuan Niu Xi	34.01	–
Niu Bang Zi	34.14	–
Tu Fu Ling	34.28	–
Suan Zao Ren	35.11	–
Fu Ling	35.45	–
(Bai) Dou Kou	35.92	–
Bai Shao Yao	36.75	–
Wu Wei Zi	36.98	–
Yuan Zhi	37.21	–
Jin Yin Hua	37.26	–
Chai Hu	37.72	–
Sha Ren	38.48	–
Bing Lang	38.69	–
Ku Shen	39.35	–
Bo He	39.54	–
Zhi Ke	39.55	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qiang Huo	40.07	–
Pu Gong Ying	40.36	–
Rou Gui	40.44	–
Xiao Hui Xiang	40.90	–
Gou Teng	40.98	–
Chen Pi	41.01	–
Zi Hua Di Ding	41.52	–
Yi Yi Ren	41.82	–
Ye Jiao Teng	42.23	–
Fo Shou	42.34	–
He Huan Pi	42.54	–
Ren Dong Teng	42.62	–
Shen Qu	42.65	–
Huang Lian	43.32	–
Ji Xue Teng	44.14	–
Dan Dou Chi	44.31	–
Ce Bai Ye	44.64	–
Chuan Xiong	44.96	–
Zhe Bei Mu	45.09	–
Du Zhong	45.25	–
Yu Jin	45.38	–
Zhu Ling	45.42	–
Cang Zhu	45.65	–
Mao Dong Qing	45.78	–
Gan Cao	45.92	–
Jing Jie	46.56	–
Sheng Jiang	47.24	–
Fu Zi	48.32	–
Gu Sui Bu	49.12	–
Xiang Fu	49.87	–
San Qi	49.90	–
Ma Huang Gen	49.94	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sang Zhi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62473	62473	0.00	9.52
62474	62474	0.00	9.35
62895	62895	0.00	8.15
62896	62896	0.00	7.57

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by

laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sha Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60184-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sha Ren; Amomi villosi fructus

Special notes

When selecting the *Sha Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sha Ren	3	0	4

Second-stage model

For differentiation of the substance/substance group *Sha Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sha Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sha Ren	G018H0939821	62401	40	from supplier
PhytoComm	Sha Ren	G018H0939821	62404	40	from supplier
PhytoComm	Sha Ren	G018H0939922	62763	40	from supplier
PhytoComm	Sha Ren	G018H0939922	62764	40	from supplier
PhytoComm	Sha Ren	G018HS179TH1	62957	40	from supplier
PhytoComm	Sha Ren	G018HS179TH1	62958	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Sha Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Sha Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sha Ren	G018H0939821	62401†	20
PhytoComm	Sha Ren	G018H0939821	62404†	20
PhytoComm	Sha Ren	G018H0939922	62763†	20
PhytoComm	Sha Ren	G018H0939922	62764†	20
PhytoComm	Sha Ren	G018HS179TH1	62957†	20
PhytoComm	Sha Ren	G018HS179TH1	62958†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 5 *Apo-Ident* customers from 5 batches from the substance/substance group *Sha Ren*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Euro OTC	Sha Ren	H0939022	1
Phytocomm	Sha Ren	g018h0939321	1
Phytocomm	Sha Ren	G018H0939321	1
Phytocomm	Sha Ren	G018H0939522	1
PhytoComm	Sha Ren	G018H0939523	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sha Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sha Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	2	120	0	12 355
Type C	1	0	5	851

The substance/substance group *Sha Ren* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	99.9714 % (> 99.9415 %)	100.0000 % (> 95.0000 %)
Type C	99.7674 % (> 99.1806 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sha Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Jiang Cao	7.97	–
Bo He	8.37	–
Qing Hao	9.54	–
Yu Jin	10.30	–
Chai Hu	10.30	–
Du Zhong	10.73	–
Dan Shen	11.85	–
Fu Zi	12.50	–
Dan Zhu Ye	12.59	–
Qiang Huo	12.69	–
Jiao Gu Lan	12.77	–
Huang Bai	12.81	–
Zi Hua Di Ding	12.85	–
Mao Dong Qing	13.64	–
Jing Jie	13.86	–
Huang Lian	14.77	–
Zhi Ke	14.82	–
Gu Sui Bu	14.95	–
Pu Gong Ying	14.97	–
Yin Chen Hao	15.94	–
Hou Po	16.64	–
Dan Dou Chi	16.98	–
Ye Jiao Teng	17.49	–
Tu Si Zi	17.72	–
Yin Yang Huo	17.76	–
Tu Fu Ling	18.46	–
He Shou Wu	18.49	–
Chen Pi	18.54	–
Fu Pen Zi	18.55	–
Hong Jing Tian	18.62	–
Ce Bai Ye	18.95	–
Shan Yao	19.21	–
Jiang Huang	19.26	–
Che Qian Zi	19.33	–
Qing Pi	19.61	–
Ze Lan	19.84	–
Bai Xian Pi	20.20	–
Bu Gu Zhi	20.22	–
(Bai) Dou Kou	20.74	–
Wu Zhu Yu	21.27	–
Sang Ye	21.49	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pi Pa Ye	21.50	–
Sang Bai Pi	22.39	–
Ling Zhi	22.43	–
Yan Hu Suo	22.55	–
Sang Ji Shend	22.75	–
Ma Huang	22.90	–
Xi Xian Cao	23.48	–
Chuan Xiong	23.63	–
Guang Huo Xiang	23.76	–
Jin Yin Hua	23.88	–
Wu Wei Zi	24.09	–
Xuan Fu Hua	25.00	–
Ren Dong Teng	25.31	–
Ji Xue Teng	25.33	–
He Huan Pi	25.94	–
Di Gu Pi	26.47	–
Yi Mu Cao	26.85	–
Ji Li	27.27	–
Mang Xiao	27.51	–
Bai Hua She She Cao	27.84	–
Tian Hua Fen	28.25	–
Shen Qu	28.42	–
Lian Zi	28.69	–
Bai Shao Yao	28.79	–
Suan Zao Ren	29.01	–
Yu Xing Cao	30.67	–
Nü Zhen Zi	30.77	–
Gan Cao	31.08	–
Chuang Mu Xiang	31.83	–
Zhu Ling	32.04	–
(Fen) Bi Xie	32.29	–
Lian Qiao	32.34	–
Ren Shen	32.38	–
(Shi) Chang Pu	32.38	–
She Gan	32.74	–
Ban Lan Gen	32.97	–
Zhi Shi	33.34	–
Xiao Hui Xiang	33.50	–
Mu Zei	33.84	–
Jin Qian Cao	34.05	–
Lu Gen	34.40	–
Gou Teng	34.52	–
Wu Yao	34.82	–
Hong Hua	34.97	–
Gan Jiang	35.09	–
Zhe Bei Mu	35.10	–
Gou Qi Zi	36.58	–
Cang Zhu	36.70	–
Cang Er Zi	37.08	–
Wang Bu Liu Xing	37.12	–
Tao Ren	37.44	–
Sheng Jiang	37.69	–
E Zhu	37.81	–
Ban Zhi Lian	38.10	–
Xin Yi	38.45	–
Wu Mei	38.79	–
Yuan Zhi	38.93	–
Chi Shao (Yao)	39.46	–
Ge Gen	39.61	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jie Geng	39.86	–
Gua Lou	40.02	–
Mu Dan Pi	40.27	–
Zhi Gan Cao	40.53	–
Huo Ma Ren	41.47	–
Shan Yu Rou	41.70	–
Hu Zhang	41.97	–
Sha Shen (Bei)	42.13	–
Chuan Niu Xi	42.30	–
Xiang Fu	42.50	–
Tai Zi Shen	42.84	–
Huang Qin	43.32	–
Du Huo	43.64	–
Rou Gui	43.85	–
Niu Bang Zi	44.09	–
Ku Shen	44.39	–
Wu Jia Pi	44.43	–
Ma Huang Gen	44.61	–
Zi Su Zi	45.29	–
Dang Gui	46.53	–
Sang Zhi	47.39	–
Fu Ling	48.01	–
Xie Bai	48.54	–
Bai Zhi	49.23	–
Hua Shi	49.40	–
Mi Huan Jun	49.41	–
Yi Yi Ren	49.77	–
Han Lian Cao	50.64	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sha Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62401	62401	0.00	19.26
62404	62404	0.00	20.60
62763	62763	0.00	12.89
62764	62764	0.00	12.59
62957	62957	0.00	8.37
62958	62958	0.00	7.97

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances,

thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sha Shen (Bei)**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60200-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sha Shen (Bei); Glehniae radix

Special notes

When selecting the *Sha Shen (Bei)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sha Shen (Bei)	2	0	1

Second-stage model

For differentiation of the substance/substance group *Sha Shen (Bei)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sha Shen (Bei)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sha Shen (Bei)	G117H0754821	62587	40	from supplier
PhytoComm	Sha Shen (Bei)	G117H0754921	62891	40	from supplier
PhytoComm	Sha Shen (Bei)	G117H0754921	62892	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 120 spectra of 3 reference samples from the substance/substance group *Sha Shen (Bei)*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 480 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 119 spectra of 4 reference samples from the substance/substance group *Sha Shen (Bei)*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sha Shen (Bei)	G117H0754821	62587†	20
PhytoComm	Sha Shen (Bei)	G117H0754821	62588	59
PhytoComm	Sha Shen (Bei)	G117H0754921	62891†	20
PhytoComm	Sha Shen (Bei)	G117H0754921	62892†	20

- 12 358 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the

substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 1 batches from the substance/substance group *Sha Shen (Bei)*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Sha Shen (Bei)	G117H0754322	4

- 853 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sha Shen (Bei)* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sha Shen (Bei)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	118	2	24 480
Type B	0	117	2	12 358
Type C	0	0	4	853

The substance/substance group *Sha Shen (Bei)* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	98.3333 % (> 95.8333 %)
Type B	100.0000 % (> 99.9402 %)	98.3193 % (> 95.7983 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sha Shen (Bei)* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Tian Hua Fen	5.22	–
Chuan Lian Zi	7.54	–
Di Gu Pi	7.96	–
Ji Li	10.06	–
E Zhu	10.13	–
Sang Zhi	11.18	–
Yan Hu Suo	13.20	–
Gan Jiang	14.54	–
Bai He	15.17	–
Jie Geng	16.21	–
Jiang Huang	17.27	–
Lai Fu Zi	17.63	–
Zi Su Zi	18.11	–
Bai Zhu	18.20	–
(Bai) Dou Kou	18.28	–
Mi Huan Jun	19.51	–
Cang Er Zi	20.05	–
Mai Ya	21.21	–
Chuan Mu Tong	22.73	–
Sha Ren	23.46	–
Dang Gui	23.60	–
Long Dan (Cao)	23.71	–
Qin Jiao	24.38	–
Mang Xiao	24.48	–
Niu Bang Zi	24.83	–
(Huai) Niu Xi	25.29	–
Huang Qi	25.53	–
(Shi) Chang Pu	25.76	–
Zhi Gan Cao	25.92	–
Shan Yao	26.28	–
He Huan Pi	27.10	–
Gua Lou	28.31	–
Ban Zhi Lian	28.49	–
Chen Pi	31.88	–
Fu Ling	32.49	–
Lian Zi	32.65	–
Bai Xian Pi	33.42	–
Xiang Fu	33.52	–
Chi Shao (Yao)	34.64	–
Chuan Niu Xi	35.40	–
Bai Hua She She Cao	36.01	–
Huang Lian	36.30	–
Chai Hu	36.85	–
Ren Dong Teng	37.05	–
Bai Zhi	37.40	–
Xiao Hui Xiang	37.52	–
Wu Wei Zi	37.57	–
Yuan Zhi	37.87	–
Bo He	37.93	–
Du Huo	38.14	–
Rou Gui	38.35	–
Tu Fu Ling	38.48	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Qiao	38.59	–
Mu Gua	38.67	–
Qiang Huo	38.72	–
Ku Shen	38.82	–
Suan Zao Ren	39.09	–
Bai Shao Yao	39.29	–
Yi Yi Ren	39.29	–
Da Zao	40.35	–
Jin Yin Hua	40.57	–
Ye Jiao Teng	42.35	–
Zhu Ling	42.48	–
Pu Gong Ying	42.65	–
Chuan Xiong	42.75	–
Shan Yu Rou	42.79	–
Ce Bai Ye	42.87	–
Shen Qu	42.92	–
Zi Hua Di Ding	43.21	–
Du Zhong	43.67	–
Gou Teng	43.74	–
Dan Dou Chi	43.87	–
Zhi Ke	43.90	–
Fo Shou	44.45	–
Yu Jin	44.80	–
Xin Yi	45.10	–
San Qi	45.57	–
Ma Huang Gen	45.74	–
Sheng Jiang	46.12	–
Cang Zhu	46.33	–
Zhu Ru	46.39	–
Mao Dong Qing	47.31	–
Xi Xian Cao	47.45	–
Wu Yao	47.48	–
Ji Xue Teng	47.61	–
Wang Bu Liu Xing	48.06	–
Sang Ji Shend	48.25	–
Gan Cao	48.67	–
Zhe Bei Mu	49.63	–
Hong Jing Tian	49.96	–
Pi Pa Ye	50.14	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sha Shen (Bei)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62587	62587	0.00	7.96
62891	62891	0.00	5.40
62892	62892	0.00	5.22

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Shan Yao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60207-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Shan Yao; Dioscoreae oppositae rhizoma

Special notes

When selecting the *Shan Yao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Shan Yao	3	0	1

Second-stage model

For differentiation of the substance/substance group *Shan Yao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Shan Yao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Shan Yao	G091H0330821	62331	40	from supplier
PhytoComm	Shan Yao	G091H0330821	62332	40	from supplier
PhytoComm	Shan Yao	G091HS024RQ1	62447	40	from supplier
PhytoComm	Shan Yao	G091HS024RQ1	62448	40	from supplier
PhytoComm	Shan Yao	G091H0330822	62523	40	from supplier
PhytoComm	Shan Yao	G091H0330822	62524	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Shan Yao*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Shan Yao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Shan Yao	G091H0330821	62331†	20
PhytoComm	Shan Yao	G091H0330821	62332†	20
PhytoComm	Shan Yao	G091HS024RQ1	62447†	20
PhytoComm	Shan Yao	G091HS024RQ1	62448†	20
PhytoComm	Shan Yao	G091H0330822	62523†	20
PhytoComm	Shan Yao	G091H0330822	62524†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 1 batches from the substance/substance group *Shan Yao*.

- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Shan Yao	G091H0330521	2
PhytoComm	Shan Yao	G091H0330521	2

- 853 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Shan Yao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Shan Yao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	239	1	24 357
Type B	13	117	3	12 344
Type C	3	4	0	850

The substance/substance group *Shan Yao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9821 % (> 99.9671 %)	99.5833 % (> 98.3333 %)
Type B	99.8500 % (> 99.8201 %)	97.5000 % (> 95.0000 %)
Type C	99.6512 % (> 99.0646 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Shan Yao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Pi Pa Ye	3.54	–
Fu Zi	4.49	–
Lian Qiao	4.73	–
Zhe Bei Mu	5.57	–
Ze Xie	6.55	–
Ling Zhi	7.47	–
Ren Dong Teng	7.90	–
Ji Li	7.93	–
Tian Hua Fen	8.20	–
Wu Wei Zi	9.52	–
Shen Qu	10.06	–
Mu Zei	10.13	–
Bai Zi Ren	10.14	–
Huo Ma Ren	10.21	–
Gua Lou	10.38	–
Chuan Xiong	10.84	–
Gu Sui Bu	11.41	–
Lian Zi	11.53	–
He Huan Pi	11.99	–
Hou Po	12.09	–
Jiao Gu Lan	12.17	–
Bai Shao Yao	12.93	–
Jin Yin Hua	12.96	–
Fu Pen Zi	13.24	–
Tai Zi Shen	13.33	–
Suan Zao Ren	13.64	–
Sheng Jiang	14.45	–
Cang Zhu	15.21	–
Fu Ling	15.25	–
Tu Fu Ling	15.40	–
Yuan Zhi	15.41	–
Chai Hu	15.78	–
Dan Dou Chi	15.86	–
Che Qian Zi	16.03	–
Yan Hu Suo	16.07	–
Ji Xue Teng	16.19	–
Chen Pi	16.37	–
Lai Fu Zi	16.37	–
Guang Huo Xiang	16.69	–
Yu Jin	16.73	–
Dan Shen	16.78	–
Ce Bai Ye	17.09	–
She Gan	17.28	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lu Gen	17.30	–
Ban Lan Gen	17.36	–
San Qi	17.50	–
Di Gu Pi	18.42	–
Mao Dong Qing	18.42	–
Gui Zhi	18.66	–
Tao Ren	19.12	–
Zhu Ling	19.37	–
Cang Er Zi	19.53	–
Ye Jiao Teng	19.68	–
Bai Xian Pi	19.80	–
Gou Teng	20.02	–
Zhu Ru	20.08	–
Yu Zhu	20.12	–
Ren Shen	20.24	–
Yin Yang Huo	20.80	–
Shan Yu Rou	21.05	–
Xie Bai	21.13	–
Dang Gui	21.28	–
Jie Geng	21.67	–
Ku Shen	21.87	–
Ci Wu Jia	22.08	–
Gan Cao	22.53	–
Rou Gui	22.67	–
Dang Gui Wei	22.91	–
Qiang Huo	22.95	–
Zhi Gan Cao	23.34	–
Hong Jing Tian	23.37	–
Zhi Ke	23.60	–
(Fen) Bi Xie	24.06	–
Gou Qi Zi	24.80	–
Ban Xia (Jiang)	24.86	–
Ma Huang	25.22	–
Qing Pi	25.78	–
Yi Yi Ren	26.01	–
Fu Xiao Mai	26.02	–
Hong Hua	26.94	–
Sang Ye	27.80	–
Qing Hao	28.55	–
Zi Hua Di Ding	29.04	–
Zhi Mu	29.63	–
Huang Qin	29.77	–
Wu Mei	29.86	–
Ma Huang Gen	30.28	–
Chuang Mu Xiang	30.35	–
Fo Shou	31.50	–
Huang Lian	31.98	–
Du Zhong	32.22	–
Jing Jie	32.34	–
Mu Gua	32.76	–
Mang Xiao	33.10	–
Bo He	33.42	–
Sha Ren	35.71	–
Sang Zhi	36.56	–
Chi Shao (Yao)	37.15	–
Long Yan Rou	37.45	–
Huang Bai	37.51	–
Fu Shen	42.68	–
Xi Xian Cao	44.71	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Jiang Cao	47.55	–
Niu Bang Zi	47.79	–
Nü Zhen Zi	48.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Shan Yao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62331	62331	0.00	11.98
62332	62332	0.00	12.08
62447	62447	0.00	3.54
62448	62448	0.00	3.61
62523	62523	0.00	10.04
62524	62524	0.00	9.52

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Shan Yu Rou**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60226-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Shan Yu Rou; Corni officinalis fructus

Special notes

When selecting the *Shan Yu Rou* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Shan Yu Rou	2	0	3

Second-stage model

For differentiation of the substance/substance group *Shan Yu Rou* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Shan Yu Rou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Shan Yu Rou	G080HS027SK1	62571	40	from supplier
PhytoComm	Shan Yu Rou	G080HS027SK1	62572	40	from supplier
PhytoComm	Shan Yu Rou	G080H0333921	63015	40	from supplier
PhytoComm	Shan Yu Rou	G080H0333921	63016	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Shan Yu Rou*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Shan Yu Rou*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Shan Yu Rou	G080HS027SK1	62571†	20
PhytoComm	Shan Yu Rou	G080HS027SK1	62572†	20
PhytoComm	Shan Yu Rou	G080H0333921	63015†	20
PhytoComm	Shan Yu Rou	G080H0333921	63016†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Shan Yu Rou*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Shan Yu Rou	g080h0333221	1
Phytocomm	Shan Yu Rou	G080H0333221	1
Phytocomm	Shan Yu Rou	G080H0333421	1
PhytoComm	Shan Yu Rou	G080H0333621	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Shan Yu Rou* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Shan Yu Rou* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	79	1	12 397
Type C	0	1	3	853

The substance/substance group *Shan Yu Rou* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Shan Yu Rou* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
(Shi) Chang Pu	6.88	–
Mu Gua	10.07	–
Yuan Zhi	12.45	–
Chen Pi	12.53	–
Wu Yao	13.15	–
Zhi Mu	13.55	–
Cang Zhu	13.84	–
Ren Shen	13.93	–
Jiao Gu Lan	15.32	–
Shan Yao	16.61	–
Bai Shao Yao	16.75	–
San Qi	16.96	–
Ku Shen	17.72	–
Pi Pa Ye	17.96	–
Jin Yin Hua	18.22	–
Chuan Lian Zi	18.50	–
Ju Hua	18.52	–
Lian Qiao	18.58	–
Gou Qi Zi	18.99	–
Mu Dan Pi	19.01	–
Dang Gui	19.60	–
Chuang Mu Xiang	19.96	–
Wu Wei Zi	20.10	–
Zhe Bei Mu	20.33	–
Xiang Fu	21.18	–
Bai Zi Ren	21.35	–
Xie Bai	21.59	–
Dang Gui Wei	21.79	–
Chuan Niu Xi	21.85	–
Jie Geng	21.92	–
Zhi Ke	22.56	–
Fang Feng	22.63	–
Sha Ren	23.14	–
Gua Lou	23.31	–
Suan Zao Ren	23.31	–
Hong Jing Tian	23.67	–
Di Gu Pi	24.59	–
He Huan Pi	24.77	–
Ji Li	24.93	–
Zhu Ru	25.09	–
Gan Cao	25.22	–
Gu Sui Bu	25.60	–
Tian Hua Fen	25.89	–
Fu Ling	26.10	–
Ren Dong Teng	26.70	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Zhu	26.97	–
Ye Jiao Teng	27.02	–
Jiang Huang	27.06	–
Xu Duan	27.42	–
Xiao Hui Xiang	27.66	–
Mang Xiao	27.87	–
Mao Dong Qing	27.88	–
Ba Ji Tian	28.00	–
Hou Po	28.15	–
Wu Zhu Yu	28.19	–
Mu Zei	28.38	–
Ling Zhi	28.89	–
Tu Fu Ling	29.51	–
Lai Fu Zi	29.75	–
Fu Zi	30.02	–
Yan Hu Suo	30.37	–
Huo Ma Ren	30.57	–
Chuan Xiong	30.83	–
Huang Qi	30.87	–
Tai Zi Shen	31.09	–
Bai Zhi	31.54	–
Sheng Jiang	31.74	–
Dan Shen	31.81	–
Lian Zi	31.94	–
Ban Lan Gen	32.10	–
Zhi Gan Cao	32.34	–
Du Huo	33.03	–
Yi Mu Cao	33.05	–
(Huai) Niu Xi	33.18	–
Sang Ji Shend	33.20	–
Fo Shou	33.23	–
Dan Dou Chi	33.66	–
Huang Qin	33.68	–
Ji Xue Teng	34.12	–
Sang Zhi	34.60	–
Mai Men Dong	35.08	–
Lu Gen	35.20	–
Sang Bai Pi	35.22	–
Niu Bang Zi	35.58	–
Fu Xiao Mai	35.72	–
Yin Chen Hao	35.77	–
Long Yan Rou	35.87	–
(Bai) Dou Kou	35.94	–
Fu Pen Zi	35.95	–
Chi Shao (Yao)	36.01	–
Zi Su Zi	36.41	–
Tao Ren	36.55	–
Ci Wu Jia	36.59	–
Bai Hua She She Cao	36.84	–
Shen Qu	36.92	–
Ban Xia (Jiang)	36.93	–
He Shou Wu	37.07	–
Gui Zhi	38.18	–
Che Qian Zi	38.52	–
Gan Jiang	38.60	–
Qing Pi	38.67	–
Chai Hu	38.89	–
Yi Yi Ren	39.28	–
Hong Hua	39.43	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Teng	39.86	–
Ze Xie	39.96	–
E Zhu	40.00	–
Rou Gui	40.71	–
Bai Zhu	41.86	–
Cang Er Zi	42.10	–
Yin Yang Huo	42.37	–
Da Zao	42.59	–
Ge Gen	43.16	–
Zi Hua Di Ding	43.40	–
Qiang Huo	43.49	–
Bing Lang	43.74	–
Mi Huan Jun	44.19	–
Bai Xian Pi	44.36	–
Xin Yi	44.99	–
Tu Si Zi	45.26	–
Huang Lian	45.79	–
Ma Huang	45.94	–
Ce Bai Ye	46.80	–
Pu Gong Ying	46.86	–
Qin Jiao	47.51	–
Sang Ye	47.53	–
Zhu Ling	47.83	–
Nü Zhen Zi	48.08	–
Du Zhong	48.56	–
Yu Jin	49.06	–
Dan Zhu Ye	49.11	–
Bo He	49.28	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Shan Yu Rou* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62571	62571	0.00	11.28
62572	62572	0.00	10.07
63015	63015	0.00	6.88
63016	63016	0.00	7.47

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**She Gan**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10003577-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

She Gan; Belamcandae rhizoma

Special notes

When selecting the *She Gan* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
She Gan	1	0	1

Second-stage model

For differentiation of the substance/substance group *She Gan* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *She Gan*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	She Gan	G048HS235TL1	62923	40	from supplier
PhytoComm	She Gan	G048HS235TL1	62924	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *She Gan*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *She Gan*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	She Gan	G048HS235TL1	62923†	20
PhytoComm	She Gan	G048HS235TL1	62924†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *She Gan*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	She Gan	G048H1034422	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *She Gan* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *She Gan* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	78	2	24 520
Type B	1	39	1	12 436
Type C	0	0	1	856

The substance/substance group *She Gan* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	97.5000 % (> 93.7500 %)
Type B	99.9857 % (> 99.9560 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *She Gan* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ren Dong Teng	3.79	–
Shen Qu	6.09	–
(Fen) Bi Xie	7.10	–
Dan Dou Chi	8.06	–
Bai Xian Pi	8.20	–
Gu Sui Bu	8.20	–
Lu Gen	8.56	–
Ce Bai Ye	8.95	–
Ling Zhi	8.98	–
Huo Ma Ren	9.59	–
Ye Jiao Teng	10.11	–
Zhu Ling	10.22	–
Guang Huo Xiang	10.28	–
Lian Zi	10.48	–
Mu Zei	10.66	–
He Huan Pi	10.88	–
Gou Teng	11.71	–
Chuan Xiong	11.85	–
Ji Xue Teng	11.97	–
Yu Jin	12.08	–
Yin Yang Huo	12.41	–
Dan Shen	12.41	–
Bai Shao Yao	13.09	–
Suan Zao Ren	13.11	–
Fu Ling	13.36	–
Pi Pa Ye	13.48	–
Shan Yao	13.70	–
Fu Pen Zi	13.80	–
Tu Fu Ling	14.26	–
Fu Zi	14.27	–
Yan Hu Suo	14.30	–
Rou Gui	14.46	–
Ji Li	14.66	–
Ma Huang Gen	14.81	–
Gui Zhi	15.23	–
Lian Qiao	15.45	–
Hou Po	15.57	–
Chai Hu	15.63	–
Tao Ren	16.01	–
Tian Hua Fen	16.22	–
Mao Dong Qing	16.46	–
Sheng Jiang	16.53	–
Zhe Bei Mu	17.27	–
Hong Jing Tian	17.30	–
Gan Cao	17.39	–
Tai Zi Shen	17.41	–
Gua Lou	17.50	–
Ma Huang	17.57	–
Che Qian Zi	17.91	–
Jin Yin Hua	18.35	–
Qiang Huo	19.70	–
Ban Lan Gen	19.89	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jing Jie	20.16	–
Yi Yi Ren	20.25	–
Ci Wu Jia	20.27	–
Jiao Gu Lan	20.79	–
Fo Shou	21.03	–
Bai Zi Ren	21.80	–
Ze Xie	22.00	–
Di Gu Pi	22.12	–
Chen Pi	22.72	–
Ban Xia (Jiang)	22.94	–
Zi Hua Di Ding	23.63	–
Cang Zhu	24.04	–
Yuan Zhi	24.15	–
Fu Xiao Mai	25.09	–
Zhi Ke	25.36	–
Yu Zhu	25.73	–
Lai Fu Zi	26.29	–
Sha Ren	26.90	–
Zhu Ru	26.99	–
Zhi Gan Cao	27.34	–
Ren Shen	28.15	–
Dang Gui Wei	29.32	–
Huang Lian	29.78	–
Bo He	29.92	–
Wu Wei Zi	30.20	–
Fu Shen	30.21	–
Jie Geng	30.42	–
Qing Pi	31.01	–
Dang Gui	31.10	–
Du Zhong	31.34	–
Chuang Mu Xiang	31.66	–
Gou Qi Zi	32.10	–
Mang Xiao	32.10	–
Huang Bai	32.39	–
San Qi	33.28	–
Huang Qin	33.50	–
Sang Zhi	35.98	–
Shan Yu Rou	38.63	–
Chi Shao (Yao)	38.96	–
Long Yan Rou	38.96	–
Xie Bai	40.22	–
Zhi Mu	40.54	–
Ku Shen	42.10	–
Bai Jiang Cao	44.11	–
Hong Hua	44.19	–
Mu Gua	44.95	–
Qing Hao	47.00	–
Cang Er Zi	49.55	–
Bai Zhu	50.01	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *She Gan* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62923	62923	0.00	4.96
62924	62924	0.00	3.79

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Shen Qu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60020-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Shen Qu; Massa fermentata medicinalis

Special notes

When selecting the *Shen Qu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Shen Qu	2	0	4

Second-stage model

For differentiation of the substance/substance group *Shen Qu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Shen Qu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Shen Qu	G156HS222RP1	62459	40	from supplier
PhytoComm	Shen Qu	G156HS222RP1	62460	40	from supplier
PhytoComm	Shen Qu	G156H0935821	62467	40	from supplier
PhytoComm	Shen Qu	G156H0935821	62468	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Shen Qu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Shen Qu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Shen Qu	G156HS222RP1	62459†	20
PhytoComm	Shen Qu	G156HS222RP1	62460†	20
PhytoComm	Shen Qu	G156H0935821	62467†	20
PhytoComm	Shen Qu	G156H0935821	62468†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Shen Qu*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Shen Qu	G156H0935121	1
phytocomm	Shen Qu	g156h0935421	1
Phytocomm	Shen Qu	G156H0935521	1
PhytoComm	Shen Qu	G156H0935521	1
PhytoComm	Shen Qu	G156H0935522	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Shen Qu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Shen Qu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	1	158	2	24 439
Type B	6	79	1	12 391
Type C	0	0	5	852

The substance/substance group *Shen Qu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	98.7500 % (> 96.8750 %)
Type B	99.9143 % (> 99.8844 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Shen Qu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Huo Ma Ren	4.32	–
Mu Zei	5.55	–
Ren Dong Teng	6.31	–
Gu Sui Bu	6.69	–
He Huan Pi	7.23	–
Lian Zi	8.28	–
Ji Li	8.56	–
Tian Hua Fen	8.77	–
Shan Yao	8.88	–
Ling Zhi	8.98	–
Ze Xie	9.46	–
Pi Pa Ye	9.58	–
Tai Zi Shen	9.65	–
Bai Shao Yao	9.68	–
Zhe Bei Mu	10.31	–
Tao Ren	10.49	–
Lian Qiao	10.54	–
Suan Zao Ren	10.92	–
Sheng Jiang	11.00	–
Fu Zi	11.25	–
Lu Gen	11.34	–
Chuan Xiong	11.36	–
Bai Zi Ren	11.51	–
Gua Lou	11.80	–
Ji Xue Teng	12.01	–
Ye Jiao Teng	12.19	–
Fu Ling	12.24	–
Gou Teng	12.57	–
She Gan	13.54	–
Lai Fu Zi	13.77	–
Ci Wu Jia	13.85	–
Ce Bai Ye	13.88	–
Bai Xian Pi	13.98	–
Gui Zhi	13.98	–
Fu Pen Zi	14.53	–
Di Gu Pi	14.64	–
Yu Jin	14.84	–
Zhu Ling	15.48	–
(Fen) Bi Xie	16.02	–
Dan Dou Chi	16.32	–
Chen Pi	16.76	–
Dan Shen	16.82	–
Yan Hu Suo	17.28	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chai Hu	17.58	–
Tu Fu Ling	17.60	–
Jin Yin Hua	17.68	–
Hou Po	18.02	–
Yuan Zhi	18.37	–
Zhu Ru	18.41	–
Rou Gui	18.46	–
Che Qian Zi	18.67	–
Ban Xia (Jiang)	18.71	–
Jiao Gu Lan	18.75	–
Ban Lan Gen	18.95	–
Guang Huo Xiang	19.06	–
Mao Dong Qing	19.57	–
Yin Yang Huo	19.61	–
Yi Yi Ren	19.75	–
Ma Huang Gen	20.39	–
Fu Xiao Mai	20.41	–
Yu Zhu	20.90	–
Wu Wei Zi	21.05	–
Fo Shou	21.33	–
Cang Zhu	21.85	–
Hong Jing Tian	22.56	–
Dang Gui Wei	22.66	–
Gan Cao	23.55	–
San Qi	23.68	–
Jie Geng	24.20	–
Ma Huang	24.20	–
Ren Shen	24.74	–
Qiang Huo	24.84	–
Dang Gui	25.39	–
Zhi Ke	25.53	–
Zhi Gan Cao	30.43	–
Gou Qi Zi	30.71	–
Jing Jie	30.72	–
Shan Yu Rou	31.58	–
Cang Er Zi	31.60	–
Huang Qin	31.62	–
Zi Hua Di Ding	31.91	–
Zhi Mu	32.11	–
Fu Shen	32.71	–
Qing Pi	32.94	–
Chuang Mu Xiang	32.95	–
Mang Xiao	33.00	–
Ku Shen	33.55	–
Huang Lian	34.22	–
Xie Bai	34.43	–
Sha Ren	35.55	–
Mu Gua	35.83	–
Bo He	36.55	–
Du Zhong	36.59	–
Long Yan Rou	36.83	–
Sang Zhi	37.01	–
Chi Shao (Yao)	40.48	–
Huang Bai	40.56	–
Yi Mu Cao	46.59	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Shen Qu* is separated from critical neighbours in a second-stage model, all

the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62459	62459	0.00	5.23
62460	62460	0.00	4.32
62467	62467	0.00	21.05
62468	62468	0.00	21.72

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Sheng Jiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60017-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Sheng Jiang; Zingiberis rhizoma recens

Special notes

When selecting the *Sheng Jiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Sheng Jiang	2	0	3

Second-stage model

For differentiation of the substance/substance group *Sheng Jiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Sheng Jiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Sheng Jiang	G253HS085SK1	62549	40	from supplier
PhytoComm	Sheng Jiang	G253HS085SK1	62550	40	from supplier
PhytoComm	Sheng Jiang	G253HS085TH2	62807	40	from supplier
PhytoComm	Sheng Jiang	G253HS085TH2	62808	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Sheng Jiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Sheng Jiang*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Sheng Jiang	G253HS085SK1	62549†	20
PhytoComm	Sheng Jiang	G253HS085SK1	62550†	20
PhytoComm	Sheng Jiang	G253HS085TH2	62807†	20
PhytoComm	Sheng Jiang	G253HS085TH2	62808†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Sheng Jiang*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Sheng Jiang	G253H0534221	1
PhytoComm	Sheng Jiang	G253H0534321	1
Phytocomm	Sheng Jiang	G253H0534522	1
PhytoComm	Sheng Jiang	G253H0534522	3

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Sheng Jiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Sheng Jiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	6	156	4	24 434
Type B	26	71	9	12 371
Type C	0	0	6	851

The substance/substance group *Sheng Jiang* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9655 % (> 99.9505 %)	97.5000 % (> 95.6250 %)
Type B	99.7512 % (> 99.7213 %)	88.7500 % (> 85.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Sheng Jiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ci Wu Jia	3.18	–
Di Gu Pi	4.05	–
Fu Xiao Mai	4.09	–
Gui Zhi	4.48	–
He Huan Pi	4.82	–
Yi Yi Ren	5.85	–
Tao Ren	5.98	–
Zhu Ru	6.15	–
Ban Xia (Jiang)	7.54	–
Lai Fu Zi	7.56	–
Fu Ling	7.58	–
Ji Li	9.07	–
Bai Zi Ren	9.21	–
Rou Gui	9.72	–
Lian Zi	10.15	–
Mu Zei	10.43	–
Gou Teng	10.75	–
Shen Qu	11.23	–
Bai Shao Yao	11.27	–
Tai Zi Shen	11.33	–
Zhe Bei Mu	11.72	–
Huo Ma Ren	12.17	–
Ling Zhi	12.83	–
Gu Sui Bu	15.11	–
Bai Xian Pi	15.24	–
Tian Hua Fen	15.65	–
Lu Gen	16.56	–
Ji Xue Teng	16.71	–
Ye Jiao Teng	16.82	–
Fo Shou	16.83	–
Gua Lou	16.99	–
Ma Huang Gen	17.36	–
Yuan Zhi	20.71	–
Ren Dong Teng	20.79	–
Shan Yao	21.10	–
Yu Jin	21.20	–
Zhu Ling	21.41	–
Ze Xie	21.70	–
Fu Shen	21.81	–
(Fen) Bi Xie	21.85	–
Pi Pa Ye	22.43	–
Tu Fu Ling	22.72	–
Chuan Xiong	22.75	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Qiao	23.06	–
Suan Zao Ren	23.71	–
Chen Pi	24.56	–
Yan Hu Suo	25.82	–
Ce Bai Ye	26.65	–
Jin Yin Hua	28.24	–
She Gan	28.42	–
Dan Shen	29.19	–
Dan Dou Chi	29.49	–
Fu Zi	30.36	–
Dang Gui Wei	30.71	–
Ren Shen	30.88	–
Jiao Gu Lan	30.90	–
Jie Geng	31.12	–
Fu Pen Zi	31.29	–
Cang Zhu	31.44	–
Mao Dong Qing	32.14	–
Chai Hu	32.51	–
Zhi Mu	32.51	–
Yin Yang Huo	32.82	–
Yu Zhu	33.12	–
San Qi	33.33	–
Shan Yu Rou	33.61	–
Mang Xiao	33.82	–
Dang Gui	34.32	–
Ban Lan Gen	34.59	–
Zhi Ke	34.87	–
Huang Qin	36.31	–
Hou Po	36.82	–
Che Qian Zi	38.13	–
Hong Jing Tian	38.66	–
Mu Gua	39.74	–
Ma Huang	40.62	–
Wu Wei Zi	40.66	–
Qiang Huo	41.04	–
Guang Huo Xiang	41.08	–
Gan Cao	41.09	–
Gou Qi Zi	42.56	–
Huang Lian	45.67	–
Long Yan Rou	46.03	–
Sang Zhi	46.34	–
Ku Shen	47.20	–
Jing Jie	48.55	–
Zi Hua Di Ding	49.85	–
Chuang Mu Xiang	50.07	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Sheng Jiang* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62549	62549	0.00	4.20
62550	62550	0.00	4.05
62807	62807	0.00	3.80
62808	62808	0.00	3.18

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Shi Gao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50368-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Shi Gao; Gypsum fibrosum

Special notes

When selecting the *Shi Gao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Shi Gao	1	0	1

Second-stage model

For differentiation of the substance/substance group *Shi Gao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Shi Gao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Shi Gao	G120HS081SH1	62561	40	from supplier
PhytoComm	Shi Gao	G120HS081SH1	62562	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Shi Gao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Shi Gao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Shi Gao	G120HS081SH1	62561†	20
PhytoComm	Shi Gao	G120HS081SH1	62562†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 1 batches from the substance/substance group *Shi Gao*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Shi Gao	G120H0541223	1
PhytoComm	Shi Gao	G120H0541223	2

- 854 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Shi Gao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Shi Gao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	3	854

The substance/substance group *Shi Gao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Shi Gao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hua Shi	53.43	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Shi Gao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62561	62561	0.00	53.91
62562	62562	0.00	53.43

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Shu Di (Huang)**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60006-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Shu Di (Huang); Rehmanniae radix praep.

Special notes

When selecting the *Shu Di (Huang)* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Shu Di (Huang)	1	0	4

Second-stage model

For differentiation of the substance/substance group *Shu Di (Huang)* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Shu Di (Huang)*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Shu Di (Huang)	G210H1601921	62989	40	from supplier
PhytoComm	Shu Di (Huang)	G210H1601921	62990	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Shu Di (Huang)*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Shu Di (Huang)*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Shu Di (Huang)	G210H1601921	62989†	20
PhytoComm	Shu Di (Huang)	G210H1601921	62990†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 15 spectra from 6 *Apo-Ident* customers from 5 batches from the substance/substance group *Shu Di (Huang)*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Hedinger	Shu Di (Huang)	g210h1601221	4
PhytoComm	Shu Di (Huang)	G210H1601121	1
Phytocomm	Shu Di (Huang)	G210H1601221	1
Phytocomm	Shu Di (Huang)	G210H1601421	6
Phytocomm	Shu Di (Huang)	G210H1601521	3

- 842 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Shu Di (Huang)* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Shu Di (Huang)* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	4	1	14	838

The substance/substance group *Shu Di (Huang)* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.6705 % (> 99.0828 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Shu Di (Huang)* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
(Sheng) Di Huang	13.39	–
Mang Xiao	16.97	–
Yu Xing Cao	23.44	–
Guang Huo Xiang	28.22	–
Hua Shi	44.20	–
Xian Mao	44.52	–
Ding Xiang	47.16	–
Jue Ming Zi	47.70	–
Jing Jie	48.95	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Shu Di (Huang)* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62989	62989	0.00	13.39
62990	62990	0.00	14.57

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Suan Zao Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60059-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Suan Zao Ren; Zizyphi spinosae semen

Special notes

When selecting the *Suan Zao Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Suan Zao Ren	3	0	2

Second-stage model

For differentiation of the substance/substance group *Suan Zao Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Suan Zao Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Suan Zao Ren	G254HS310SM1	62717	40	from supplier
PhytoComm	Suan Zao Ren	G254HS310SM1	62718	40	from supplier
PhytoComm	Suan Zao Ren	G254HS310SV1	62809	40	from supplier
PhytoComm	Suan Zao Ren	G254HS310SV1	62810	40	from supplier
PhytoComm	Suan Zao Ren	G254H1418021	62987	40	from supplier
PhytoComm	Suan Zao Ren	G254H1418021	62988	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Suan Zao Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Suan Zao Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Suan Zao Ren	G254HS310SM1	62717†	20
PhytoComm	Suan Zao Ren	G254HS310SM1	62718†	20
PhytoComm	Suan Zao Ren	G254HS310SV1	62809†	20
PhytoComm	Suan Zao Ren	G254HS310SV1	62810†	20
PhytoComm	Suan Zao Ren	G254H1418021	62987†	20
PhytoComm	Suan Zao Ren	G254H1418021	62988†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Suan Zao Ren*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Suan Zao Ren	g254h1418121	2
Phytocomm	Suan Zao Ren	G254H1418421	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Suan Zao Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Suan Zao Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	118	2	12 357
Type C	1	0	3	853

The substance/substance group *Suan Zao Ren* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	98.3333 % (> 95.8333 %)
Type C	99.8837 % (> 99.2977 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Suan Zao Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lian Zi	5.47	–
Bai Xian Pi	8.29	–
Lian Qiao	8.44	–
She Gan	9.19	–
Lu Gen	10.53	–
Ren Dong Teng	11.32	–
Cang Zhu	11.42	–
Ji Li	11.48	–
He Huan Pi	11.69	–
Ling Zhi	11.70	–
Bai Zi Ren	12.00	–
Tu Fu Ling	12.21	–
Zhe Bei Mu	12.24	–
Jiao Gu Lan	12.28	–
Tao Ren	12.56	–
Ye Jiao Teng	12.62	–
Dan Dou Chi	12.87	–
Shen Qu	13.04	–
Shan Yao	13.26	–
Xie Bai	13.47	–
Mao Dong Qing	13.71	–
Gu Sui Bu	13.98	–
Bai Shao Yao	14.01	–
Gua Lou	14.02	–
San Qi	14.28	–
Rou Gui	14.44	–
Gou Teng	14.54	–
Chuan Xiong	14.89	–
Mu Zei	15.10	–
Huo Ma Ren	15.16	–
Guang Huo Xiang	15.32	–
Sheng Jiang	15.59	–
Yuan Zhi	15.90	–
Ji Xue Teng	16.40	–
Gui Zhi	16.41	–
Jin Yin Hua	16.47	–
Ce Bai Ye	16.47	–
Pi Pa Ye	16.88	–
Ku Shen	16.89	–
Chai Hu	17.16	–
Tai Zi Shen	17.18	–
Fu Ling	17.30	–
Chen Pi	17.46	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
(Fen) Bi Xie	17.91	–
Dan Shen	18.10	–
Fu Zi	18.28	–
Tian Hua Fen	18.62	–
Ci Wu Jia	18.70	–
Yu Jin	19.22	–
Cang Er Zi	19.34	–
Fu Pen Zi	19.50	–
Hong Jing Tian	19.84	–
Yu Zhu	19.87	–
Di Gu Pi	20.25	–
Wu Wei Zi	20.33	–
Ma Huang Gen	20.37	–
Ban Lan Gen	20.63	–
Ze Xie	21.10	–
Hou Po	21.21	–
Gan Cao	21.42	–
Fo Shou	21.52	–
Ma Huang	21.58	–
Zhu Ling	21.64	–
Yin Yang Huo	22.38	–
Gou Qi Zi	23.11	–
Zhi Ke	23.11	–
Chuang Mu Xiang	23.33	–
Yi Yi Ren	23.82	–
Zhu Ru	23.93	–
Shan Yu Rou	24.22	–
Dang Gui Wei	24.22	–
Che Qian Zi	25.47	–
Lai Fu Zi	25.63	–
Jing Jie	26.09	–
Ban Xia (Jiang)	26.31	–
E Zhu	26.43	–
Fu Xiao Mai	26.53	–
Qiang Huo	27.01	–
Mang Xiao	27.17	–
Ren Shen	27.27	–
Dang Gui	27.59	–
Yan Hu Suo	28.04	–
Jie Geng	29.03	–
Zi Hua Di Ding	29.03	–
Zhi Gan Cao	31.42	–
Zi Su Zi	31.96	–
Jiang Huang	32.64	–
Dong Gua Zi	32.79	–
Mu Gua	32.98	–
Zhi Mu	33.04	–
Bo He	33.88	–
Qing Pi	33.93	–
Fu Shen	33.95	–
Du Zhong	34.15	–
Gan Jiang	35.87	–
Sha Ren	36.04	–
Niu Bang Zi	36.26	–
Huang Bai	36.27	–
Long Yan Rou	36.29	–
Sang Zhi	37.51	–
Huang Qin	37.63	–
Huang Lian	38.45	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Shen (Bei)	40.45	–
Hong Hua	40.86	–
Chi Shao (Yao)	41.32	–
Qing Hao	44.20	–
Chuan Lian Zi	44.93	–
Long Dan (Cao)	45.87	–
(Huai) Niu Xi	46.30	–
Bai Zhu	47.50	–
Sang Ye	49.13	–
(Bai) Dou Kou	49.47	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Suan Zao Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62717	62717	0.00	8.44
62718	62718	0.00	9.09
62809	62809	0.00	5.47
62810	62810	0.00	6.73
62987	62987	0.00	19.91
62988	62988	0.00	19.34

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Tai Zi Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60114-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Tai Zi Shen; Pseudostellariae heterophyllae radix

Special notes

When selecting the *Tai Zi Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Tai Zi Shen	1	0	2

Second-stage model

For differentiation of the substance/substance group *Tai Zi Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Tai Zi Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tai Zi Shen	G204HM066SK1	62595	40	from supplier
PhytoComm	Tai Zi Shen	G204HM066SK1	62596	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Tai Zi Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Tai Zi Shen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Tai Zi Shen	G204HM066SK1	62595†	20
PhytoComm	Tai Zi Shen	G204HM066SK1	62596†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Tai Zi Shen*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Tai Zi Shen	G204H0446221	1
Phytocomm	Tai Zi Shen	G204H0446521	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Tai Zi Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tai Zi Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	4	40	0	12 433
Type C	0	0	2	855

The substance/substance group *Tai Zi Shen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9543 % (> 99.9245 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tai Zi Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Shao Yao	5.48	–
Zhe Bei Mu	6.05	–
Shen Qu	6.42	–
Mu Zei	6.64	–
He Huan Pi	6.88	–
Gui Zhi	7.20	–
Ye Jiao Teng	7.78	–
Lai Fu Zi	7.84	–
Tian Hua Fen	7.97	–
Ban Xia (Jiang)	8.29	–
Sheng Jiang	8.47	–
Tao Ren	8.48	–
Huo Ma Ren	9.40	–
Ci Wu Jia	9.64	–
Ji Li	9.69	–
Rou Gui	10.01	–
Ji Xue Teng	10.33	–
Di Gu Pi	10.72	–
Gou Teng	10.76	–
Gua Lou	10.85	–
Fo Shou	11.04	–
Bai Zi Ren	11.69	–
Zhu Ru	12.35	–
Gu Sui Bu	12.37	–
Lu Gen	12.52	–
Pi Pa Ye	12.65	–
Lian Qiao	13.05	–
Ling Zhi	13.87	–
Jin Yin Hua	14.04	–
Ze Xie	14.07	–
Lian Zi	15.09	–
Fu Ling	15.99	–
Ren Dong Teng	16.26	–
Fu Xiao Mai	16.61	–
Shan Yao	16.99	–
Bai Xian Pi	17.33	–
Yuan Zhi	17.40	–
Fu Zi	17.66	–
Chen Pi	17.73	–
Ma Huang Gen	18.15	–
Yan Hu Suo	18.26	–
Yi Yi Ren	18.27	–
Chuan Xiong	18.99	–
Fu Pen Zi	19.60	–
Dang Gui Wei	20.02	–
Yu Zhu	20.74	–
Dang Gui	20.79	–
Ban Lan Gen	21.04	–
(Fen) Bi Xie	21.08	–
Jie Geng	21.38	–
Yu Jin	21.94	–
Zhu Ling	22.10	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	22.54	–
Tu Fu Ling	22.67	–
Ren Shen	22.70	–
Ce Bai Ye	22.75	–
Fu Shen	23.13	–
Hou Po	23.28	–
Suan Zao Ren	23.35	–
Zhi Mu	23.55	–
Zhi Ke	23.60	–
Ma Huang	23.69	–
Cang Zhu	23.76	–
Dan Shen	23.85	–
Yin Yang Huo	25.25	–
Huang Qin	25.72	–
Dan Dou Chi	26.01	–
San Qi	26.09	–
She Gan	27.96	–
Gou Qi Zi	28.04	–
Ku Shen	28.21	–
Che Qian Zi	28.28	–
Shan Yu Rou	28.49	–
Wu Wei Zi	28.67	–
Chai Hu	29.53	–
Guang Huo Xiang	29.95	–
Mao Dong Qing	30.23	–
Hong Jing Tian	31.43	–
Mu Gua	31.51	–
Mang Xiao	31.95	–
Long Yan Rou	32.40	–
Qiang Huo	32.71	–
Gan Cao	33.23	–
Huang Lian	38.66	–
Sang Zhi	38.89	–
Chuang Mu Xiang	41.36	–
Xie Bai	41.57	–
Zi Hua Di Ding	41.78	–
Qing Pi	41.92	–
Chi Shao (Yao)	42.22	–
Jing Jie	42.23	–
Sha Ren	45.00	–
Zhi Gan Cao	45.70	–
Du Zhong	47.22	–
Bo He	47.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tai Zi Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62595	62595	0.00	6.09
62596	62596	0.00	5.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Tao Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60454-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Tao Ren; Persicae semen

Special notes

When selecting the *Tao Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Tao Ren	1	0	2

Second-stage model

For differentiation of the substance/substance group *Tao Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Tao Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tao Ren	G184HS220SK1	62609	40	from supplier
PhytoComm	Tao Ren	G184HS220SK1	62610	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Tao Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Tao Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Tao Ren	G184HS220SK1	62609†	20
PhytoComm	Tao Ren	G184HS220SK1	62610†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Tao Ren*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Tao Ren	g184h1007122	1
PhytoComm	Tao Ren	G184H1007321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Tao Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tao Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	75	5	24 519
Type B	1	32	8	12 436
Type C	0	0	2	855

The substance/substance group *Tao Ren* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9779 %)	93.7500 % (> 90.0000 %)
Type B	99.9964 % (> 99.9667 %)	80.0000 % (> 72.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tao Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
He Huan Pi	3.66	–
Gui Zhi	5.80	–
Ci Wu Jia	5.88	–
Sheng Jiang	6.33	–
Lai Fu Zi	6.36	–
Mu Zei	7.92	–
Lian Zi	8.67	–
Rou Gui	9.21	–
Huo Ma Ren	9.66	–
Ling Zhi	10.00	–
Fu Xiao Mai	10.09	–
Bai Zi Ren	10.13	–
Bai Shao Yao	10.29	–
Gou Teng	10.77	–
Di Gu Pi	11.00	–
Tai Zi Shen	11.22	–
Ban Xia (Jiang)	11.39	–
Bai Xian Pi	11.72	–
Yi Yi Ren	11.92	–
Fu Ling	12.38	–
Ye Jiao Teng	12.39	–
Fo Shou	12.83	–
Ji Li	12.85	–
Gu Sui Bu	12.86	–
Zhu Ru	13.03	–
Shen Qu	13.65	–
Lu Gen	13.88	–
Ji Xue Teng	14.91	–
Zhe Bei Mu	15.12	–
Suan Zao Ren	15.94	–
Ze Xie	16.07	–
Gua Lou	16.10	–
Ma Huang Gen	16.75	–
Tu Fu Ling	17.72	–
Tian Hua Fen	17.98	–
Ren Dong Teng	18.14	–
Chuan Xiong	20.26	–
She Gan	20.37	–
(Fen) Bi Xie	20.53	–
Yuan Zhi	21.39	–
Zhu Ling	21.66	–
Yu Jin	21.67	–
Chen Pi	22.02	–
Pi Pa Ye	22.39	–
Shan Yao	22.82	–
Fu Shen	23.88	–
Lian Qiao	24.36	–
Ce Bai Ye	24.67	–
Jin Yin Hua	24.77	–
Dan Dou Chi	25.04	–
Dan Shen	27.45	–
Cang Zhu	28.22	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mao Dong Qing	28.66	–
Chai Hu	28.71	–
Yu Zhu	28.82	–
Jiao Gu Lan	29.27	–
Yan Hu Suo	29.44	–
Jie Geng	29.66	–
Fu Pen Zi	29.75	–
Ban Lan Gen	29.90	–
Dang Gui Wei	30.13	–
Fu Zi	30.14	–
San Qi	30.98	–
Hong Jing Tian	30.98	–
Guang Huo Xiang	31.12	–
Dang Gui	31.59	–
Zhi Ke	31.74	–
Yin Yang Huo	31.77	–
Zhi Mu	32.10	–
Mang Xiao	33.45	–
Ma Huang	34.71	–
Shan Yu Rou	34.85	–
Ren Shen	35.17	–
Hou Po	35.22	–
Gan Cao	36.73	–
Che Qian Zi	38.11	–
Qiang Huo	38.88	–
Mu Gua	39.38	–
Huang Qin	39.50	–
Wu Wei Zi	41.90	–
Chuang Mu Xiang	42.03	–
Jing Jie	42.26	–
Gou Qi Zi	44.16	–
Long Yan Rou	44.36	–
Zi Hua Di Ding	44.40	–
Sang Zhi	45.33	–
Huang Lian	46.17	–
Sha Ren	47.92	–
Ku Shen	48.32	–
Xie Bai	49.31	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tao Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62609	62609	0.00	3.66
62610	62610	0.00	3.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Tian Hua Fen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60139-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Tian Hua Fen; Trichosanthis radix

Special notes

When selecting the *Tian Hua Fen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Tian Hua Fen	3	0	3

Second-stage model

For differentiation of the substance/substance group *Tian Hua Fen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Tian Hua Fen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tian Hua Fen	G242H0417821	62315	40	from supplier
PhytoComm	Tian Hua Fen	G242H0417821	62316	40	from supplier
PhytoComm	Tian Hua Fen	G242HS040SG1	62633	40	from supplier
PhytoComm	Tian Hua Fen	G242HS040SG1	62634	40	from supplier
PhytoComm	Tian Hua Fen	G242HS040TK1	62799	40	from supplier
PhytoComm	Tian Hua Fen	G242HS040TK1	62800	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Tian Hua Fen*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Tian Hua Fen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Tian Hua Fen	G242H0417821	62315†	20
PhytoComm	Tian Hua Fen	G242H0417821	62316†	20
PhytoComm	Tian Hua Fen	G242HS040SG1	62633†	20
PhytoComm	Tian Hua Fen	G242HS040SG1	62634†	20
PhytoComm	Tian Hua Fen	G242HS040TK1	62799†	20
PhytoComm	Tian Hua Fen	G242HS040TK1	62800†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Tian Hua Fen*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Tian Hua Fen	G242H0417423	2
Phytocomm	Tian Hua Fen	G242H0417522	1
PhytoComm	Tian Hua Fen	G242H0417522	1
Phytocomm	Tian Hua Fen	h0417022	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Tian Hua Fen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tian Hua Fen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	2	240	0	24 358
Type B	6	119	1	12 351
Type C	0	0	5	852

The substance/substance group *Tian Hua Fen* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9905 % (> 99.9754 %)	100.0000 % (> 97.5000 %)
Type B	99.9547 % (> 99.9248 %)	99.1667 % (> 96.6667 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tian Hua Fen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lai Fu Zi	5.23	–
Mu Zei	5.57	–
Tai Zi Shen	6.20	–
He Huan Pi	6.26	–
Zhe Bei Mu	7.03	–
Huo Ma Ren	8.14	–
Shen Qu	8.44	–
Lian Qiao	9.06	–
Bai Shao Yao	9.20	–
Pi Pa Ye	9.25	–
Bai Zi Ren	9.26	–
Ji Li	9.40	–
Sha Shen (Bei)	9.76	–
Gu Sui Bu	9.83	–
Fu Zi	10.12	–
Ze Xie	10.35	–
Shan Yao	10.88	–
Gua Lou	11.05	–
Ji Xue Teng	11.33	–
Gui Zhi	11.45	–
Sheng Jiang	11.65	–
Di Gu Pi	11.69	–
Tao Ren	11.92	–
Ren Dong Teng	11.99	–
Jin Yin Hua	12.09	–
Ling Zhi	12.19	–
Bai He	12.28	–
Sang Zhi	12.86	–
Ban Xia (Jiang)	13.04	–
Ye Jiao Teng	13.34	–
Lu Gen	13.86	–
Yan Hu Suo	13.99	–
Chuan Lian Zi	14.20	–
Lian Zi	14.36	–
Rou Gui	14.44	–
Fu Pen Zi	15.12	–
Gou Teng	15.51	–
Hou Po	15.92	–
Zhu Ru	16.07	–
Yuan Zhi	16.31	–
Gan Jiang	16.46	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Shen	16.63	–
Ci Wu Jia	16.84	–
Dang Gui	16.87	–
Ku Shen	17.47	–
Chuan Xiong	17.54	–
Jie Geng	17.95	–
Yu Jin	18.04	–
Ce Bai Ye	18.06	–
Zhu Ling	18.22	–
Chen Pi	18.29	–
Yi Yi Ren	18.34	–
Bai Xian Pi	18.50	–
Fu Ling	18.55	–
Fu Xiao Mai	18.68	–
Dang Gui Wei	18.70	–
Suan Zao Ren	18.72	–
Cang Zhu	18.99	–
Jiao Gu Lan	19.36	–
(Shi) Chang Pu	19.55	–
Ban Lan Gen	19.67	–
Che Qian Zi	19.72	–
Yu Zhu	20.26	–
Zhi Ke	20.33	–
Dan Shen	20.56	–
Dan Dou Chi	20.78	–
Huang Qin	20.78	–
Ma Huang	21.07	–
(Fen) Bi Xie	21.28	–
E Zhu	21.63	–
Chuan Mu Tong	21.63	–
San Qi	22.00	–
Zhi Mu	22.51	–
Yin Yang Huo	22.64	–
Fo Shou	23.07	–
She Gan	23.08	–
Cang Er Zi	23.26	–
Zi Su Zi	23.28	–
Wu Wei Zi	23.76	–
Guang Huo Xiang	23.94	–
Ma Huang Gen	24.03	–
Chai Hu	24.16	–
Gou Qi Zi	24.21	–
Shan Yu Rou	24.23	–
Mai Ya	24.36	–
Mi Huan Jun	25.00	–
Mao Dong Qing	25.59	–
Tu Fu Ling	25.60	–
Huang Qi	25.97	–
Gan Cao	26.79	–
Zhi Gan Cao	27.06	–
Bai Zhu	27.75	–
Qiang Huo	27.81	–
Hong Jing Tian	27.97	–
Mang Xiao	28.73	–
Qin Jiao	28.75	–
Jiang Huang	28.97	–
Niu Bang Zi	29.45	–
(Huai) Niu Xi	29.56	–
(Bai) Dou Kou	30.70	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pu Gong Ying	31.09	–
Mu Gua	31.58	–
Ban Zhi Lian	32.40	–
Long Yan Rou	32.63	–
Long Dan (Cao)	32.72	–
Fu Shen	33.21	–
Xiao Hui Xiang	33.84	–
Qing Pi	33.93	–
Bo He	34.68	–
Huang Lian	34.78	–
Zi Hua Di Ding	36.30	–
Xie Bai	37.17	–
Chi Shao (Yao)	37.73	–
Chuang Mu Xiang	38.08	–
Jing Jie	38.65	–
Sha Ren	38.95	–
Bai Zhi	39.47	–
Du Zhong	39.87	–
Yi Mu Cao	41.04	–
Chuan Niu Xi	43.62	–
Huang Bai	45.70	–
Da Zao	46.32	–
Mai Men Dong	48.01	–
Xiang Fu	48.40	–
Sang Ye	49.16	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tian Hua Fen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62315	62315	0.00	10.63
62316	62316	0.00	9.76
62633	62633	0.00	5.57
62634	62634	0.00	5.23
62799	62799	0.00	7.03
62800	62800	0.00	7.13

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Tu Fu Ling**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50885-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Tu Fu Ling; Smilacis glabrae rhizoma

Special notes

When selecting the *Tu Fu Ling* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Tu Fu Ling	3	0	3

Second-stage model

For differentiation of the substance/substance group *Tu Fu Ling* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Tu Fu Ling*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tu Fu Ling	G231HS009RK1	62427	40	from supplier
PhytoComm	Tu Fu Ling	G231HS009RK1	62428	40	from supplier
PhytoComm	Tu Fu Ling	G231HS009SK1	62641	40	from supplier
PhytoComm	Tu Fu Ling	G231HS009SK1	62642	40	from supplier
PhytoComm	Tu Fu Ling	G231HS009TP1	63033	40	from supplier
PhytoComm	Tu Fu Ling	G231HS009TP1	63034	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Tu Fu Ling*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Tu Fu Ling*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Tu Fu Ling	G231HS009RK1	62427†	20
PhytoComm	Tu Fu Ling	G231HS009RK1	62428†	20
PhytoComm	Tu Fu Ling	G231HS009SK1	62641†	20
PhytoComm	Tu Fu Ling	G231HS009SK1	62642†	20
PhytoComm	Tu Fu Ling	G231HS009TP1	63033†	20
PhytoComm	Tu Fu Ling	G231HS009TP1	63034†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Tu Fu Ling*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Tu Fu Ling	G231H0325121	1
PhytoComm	Tu Fu Ling	G231H0325121	1
Phytocomm	Tu Fu Ling	G231H0325421	2
PhytoComm	Tu Fu Ling	G231H0325421	1
PhytoComm	Tu Fu Ling	G23H0325121	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Tu Fu Ling* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tu Fu Ling* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	117	3	12 357
Type C	2	0	6	849

The substance/substance group *Tu Fu Ling* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	97.5000 % (> 95.0000 %)
Type C	99.4186 % (> 98.8315 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tu Fu Ling* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ze Lan	5.41	–
Bai Xian Pi	5.56	–
Ye Jiao Teng	8.04	–
Xi Xian Cao	8.49	–
Chai Hu	9.14	–
Dan Dou Chi	9.62	–
Yin Chen Hao	11.72	–
Ji Xue Teng	12.95	–
Rou Gui	13.18	–
Mao Dong Qing	13.32	–
Pu Gong Ying	13.32	–
Ling Zhi	13.54	–
Du Zhong	13.55	–
Yin Yang Huo	14.62	–
Dan Shen	14.82	–
Lian Zi	14.92	–
Guang Huo Xiang	15.29	–
Fu Zi	15.48	–
Ren Dong Teng	16.50	–
Zhi Ke	17.40	–
Bai Jiang Cao	17.55	–
Sha Ren	17.73	–
Ce Bai Ye	17.74	–
Xin Yi	17.78	–
Huang Lian	17.98	–
(Fen) Bi Xie	18.51	–
Jiao Gu Lan	18.72	–
Shan Yao	19.11	–
Jing Jie	19.45	–
Yu Jin	19.66	–
Zhu Ling	19.89	–
Zi Hua Di Ding	19.94	–
Nü Zhen Zi	19.94	–
Gou Teng	19.95	–
Bai Shao Yao	20.18	–
Qiang Huo	20.19	–
Hou Po	20.54	–
Shen Qu	20.65	–
Che Qian Zi	20.82	–
Fo Shou	21.14	–
Lu Gen	21.19	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	21.69	–
Ma Huang Gen	21.74	–
Bo He	21.83	–
Sang Ye	22.06	–
He Huan Pi	22.12	–
Zhi Shi	23.13	–
Huang Bai	23.35	–
Gan Cao	23.64	–
Huo Ma Ren	23.83	–
Bai Hua She She Cao	24.03	–
Ma Huang	24.15	–
Gu Sui Bu	24.54	–
Di Gu Pi	25.08	–
Fu Ling	25.18	–
Gui Zhi	25.20	–
Ji Li	25.27	–
Jiang Huang	25.36	–
Sheng Jiang	25.39	–
Yu Xing Cao	25.87	–
Yi Yi Ren	25.96	–
Ban Zhi Lian	26.28	–
Yan Hu Suo	26.73	–
Pi Pa Ye	27.62	–
She Gan	27.90	–
Gua Lou	28.56	–
Mang Xiao	28.73	–
Lian Qiao	28.97	–
Fu Pen Zi	29.01	–
Chuan Xiong	29.21	–
Chuang Mu Xiang	29.72	–
Cang Er Zi	30.45	–
Suan Zao Ren	30.73	–
Hu Zhang	30.75	–
Cang Zhu	30.93	–
Tao Ren	31.37	–
Fu Shen	32.32	–
Wu Wei Zi	32.62	–
Qing Pi	32.75	–
Wu Mei	32.98	–
Zhe Bei Mu	32.99	–
Tian Hua Fen	33.05	–
Mu Zei	33.10	–
Tai Zi Shen	33.21	–
Yuan Zhi	33.28	–
Sang Zhi	33.53	–
Hong Jing Tian	34.04	–
Ban Xia (Jiang)	34.21	–
Ci Wu Jia	34.99	–
Chi Shao (Yao)	34.99	–
Fu Xiao Mai	35.18	–
Gou Qi Zi	35.25	–
Wang Bu Liu Xing	35.79	–
Zhi Gan Cao	35.91	–
Jie Geng	36.38	–
Han Lian Cao	36.41	–
Qing Hao	36.82	–
Xiang Fu	37.17	–
Ban Lan Gen	37.39	–
E Zhu	37.45	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Su Zi	37.96	–
Ren Shen	38.89	–
Chen Pi	39.15	–
Yi Mu Cao	39.37	–
Huang Qin	39.91	–
Bai Zhu	40.69	–
Zhu Ru	41.05	–
Yu Zhu	41.64	–
(Shi) Chang Pu	42.39	–
Dang Gui Wei	43.82	–
Niu Bang Zi	44.38	–
Bai Zi Ren	44.62	–
Hong Hua	44.76	–
(Bai) Dou Kou	45.18	–
Dang Gui	46.18	–
Sha Shen (Bei)	47.07	–
Mi Huan Jun	47.52	–
Long Yan Rou	47.72	–
Huang Qi	48.12	–
San Qi	48.95	–
Lai Fu Zi	49.88	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tu Fu Ling* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62427	62427	0.00	5.56
62428	62428	0.00	6.99
62641	62641	0.00	5.41
62642	62642	0.00	5.69
63033	63033	0.00	13.80
63034	63034	0.00	14.85

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Tu Si Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60217-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Tu Si Zi; Cuscutae semen

Special notes

When selecting the *Tu Si Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Tu Si Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Tu Si Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Tu Si Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tu Si Zi	G086H1243823	62521	40	from supplier
PhytoComm	Tu Si Zi	G086H1243823	62522	40	from supplier
PhytoComm	Tu Si Zi	G086H1243924	62935	40	from supplier
PhytoComm	Tu Si Zi	G086H1243924	62936	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Tu Si Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Tu Si Zi*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Tu Si Zi	G086H1243823	62521†	20
PhytoComm	Tu Si Zi	G086H1243823	62522†	20
PhytoComm	Tu Si Zi	G086H1243924	62935†	20
PhytoComm	Tu Si Zi	G086H1243924	62936†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Tu Si Zi*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Tu Si Zi	G086H1243521	3
PhytoComm	Tu Si Zi	G086H1243521	1
Phytocomm	Tu Si Zi	h1243022	1

- 852 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Tu Si Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Tu Si Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	5	852

The substance/substance group *Tu Si Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Tu Si Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bu Gu Zhi	14.64	–
Gu Sui Bu	22.32	–
Wu Yao	22.51	–
Yin Chen Hao	22.81	–
Mang Xiao	23.59	–
(Bai) Dou Kou	25.40	–
Yu Xing Cao	25.63	–
Sang Bai Pi	26.14	–
Sha Ren	28.09	–
He Huan Pi	28.49	–
Sang Ji Shend	31.21	–
Dan Zhu Ye	31.60	–
E Zhu	34.99	–
Jiang Huang	35.62	–
Ji Li	36.33	–
(Sheng) Di Huang	39.15	–
He Shou Wu	41.55	–
Xuan Fu Hua	42.80	–
(Shi) Chang Pu	44.46	–
Jing Jie	44.86	–
Shan Yao	46.12	–
Hua Shi	47.34	–
Wu Zhu Yu	48.99	–
Wu Jia Pi	49.69	–
Chen Pi	50.90	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tu Si Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62521	62521	0.00	24.46
62522	62522	0.00	24.19

continued on the next page

continued from previous page

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62935	62935	0.00	14.64
62936	62936	0.00	15.50

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wang Bu Liu Xing**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50390-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wang Bu Liu Xing; Vaccariae semen

Special notes

When selecting the *Wang Bu Liu Xing* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wang Bu Liu Xing	1	0	4

Second-stage model

For differentiation of the substance/substance group *Wang Bu Liu Xing* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wang Bu Liu Xing*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wang Bu Liu Xing	G281H0456921	62731	40	from supplier
PhytoComm	Wang Bu Liu Xing	G281H0456921	62732	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Wang Bu Liu Xing*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Wang Bu Liu Xing*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wang Bu Liu Xing	G281H0456921	62731†	20
PhytoComm	Wang Bu Liu Xing	G281H0456921	62732†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 4 batches from the substance/substance group *Wang Bu Liu Xing*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Wang Bu Liu Xing	G281H0456221	1
PhytoComm	Wang Bu Liu Xing	G281H0456322	1
Phytocomm	Wang Bu Liu Xing	G281H0456521	1
PhytoComm	Wang Bu Liu Xing	G281H456421	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wang Bu Liu Xing* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wang Bu Liu Xing* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	4	853

The substance/substance group *Wang Bu Liu Xing* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wang Bu Liu Xing* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Cang Er Zi	23.98	–
Mang Xiao	25.70	–
E Zhu	26.55	–
He Huan Pi	26.62	–
Jiang Huang	26.89	–
Zi Su Zi	29.68	–
Sha Ren	30.35	–
Ji Li	31.10	–
Shan Yao	32.21	–
Niu Bang Zi	33.38	–
Qiang Huo	36.26	–
Suan Zao Ren	36.94	–
(Bai) Dou Kou	37.49	–
Yan Hu Suo	37.79	–
Ze Lan	37.86	–
Huang Lian	38.03	–
Gan Jiang	38.34	–
Du Zhong	38.47	–
Chen Pi	38.66	–
Bai Hua She She Cao	38.93	–
Yi Mu Cao	39.26	–
Wu Mei	40.45	–
Mao Dong Qing	41.95	–
Xin Yi	42.22	–
Nü Zhen Zi	42.81	–
Yin Chen Hao	43.07	–
Sha Shen (Bei)	43.20	–
Dan Dou Chi	43.81	–
Xiang Fu	45.29	–
Xia Ku Cao	45.39	–
Xi Xian Cao	45.83	–
Bo He	46.89	–
Zhi Ke	47.12	–
Hu Zhang	47.15	–
Sang Zhi	48.25	–
Che Qian Zi	48.57	–
Bai Xian Pi	49.20	–
Di Gu Pi	50.13	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wang Bu Liu Xing* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62731	62731	0.00	25.00
62732	62732	0.00	23.98

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wu Jia Pi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002259-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wu Jia Pi; Acanthopanacis cortex

Special notes

When selecting the *Wu Jia Pi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wu Jia Pi	1	0	1

Second-stage model

For differentiation of the substance/substance group *Wu Jia Pi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wu Jia Pi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wu Jia Pi	G002H0402921	62757	40	from supplier
PhytoComm	Wu Jia Pi	G002H0402921	62758	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Wu Jia Pi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Wu Jia Pi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wu Jia Pi	G002H0402921	62757†	20
PhytoComm	Wu Jia Pi	G002H0402921	62758†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Wu Jia Pi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Wu Jia Pi	g002h0402321	1
Phytocomm	Wu Jia Pi	G002H0402321	1
PhytoComm	Wu Jia Pi	G002H0402321	3

- 852 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wu Jia Pi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wu Jia Pi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	5	852

The substance/substance group *Wu Jia Pi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wu Jia Pi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Dan Zhu Ye	15.56	–
Yu Xing Cao	19.05	–
Jing Jie	19.65	–
Xian Mao	22.01	–
Mang Xiao	22.95	–
Sang Ye	24.53	–
Sang Ji Shend	25.07	–
Du Zhong	26.63	–
Ding Xiang	28.03	–
(Sheng) Di Huang	32.76	–
Yin Chen Hao	32.95	–
Bu Gu Zhi	33.80	–
Ge Gen	35.88	–
Jin Qian Cao	35.88	–
Xuan Shen	41.30	–
Sha Ren	43.98	–
He Shou Wu	47.58	–
Hua Shi	47.86	–
Nü Zhen Zi	49.65	–
Zhi Shi	50.46	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wu Jia Pi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62757	62757	0.00	15.84
62758	62758	0.00	15.56

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wu Mei**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60021-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wu Mei; Pruni mume fructus

Special notes

When selecting the *Wu Mei* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wu Mei	2	0	2

Second-stage model

For differentiation of the substance/substance group *Wu Mei* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wu Mei*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wu Mei	G167H1026821	62477	40	from supplier
PhytoComm	Wu Mei	G167H1026821	62478	40	from supplier
PhytoComm	Wu Mei	G167HS192SK1	62715	40	from supplier
PhytoComm	Wu Mei	G167HS192SK1	62716	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Wu Mei*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Wu Mei*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wu Mei	G167H1026821	62477†	20
PhytoComm	Wu Mei	G167H1026821	62478†	20
PhytoComm	Wu Mei	G167HS192SK1	62715†	20
PhytoComm	Wu Mei	G167HS192SK1	62716†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Wu Mei*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Wu Mei	G167H1026221	1
Phytocomm	Wu Mei	G167H1026421	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wu Mei* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wu Mei* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	2	855

The substance/substance group *Wu Mei* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wu Mei* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sang Ye	14.91	–
Shan Yao	17.71	–
Qing Pi	18.49	–
Huang Bai	23.42	–
Mang Xiao	24.29	–
Qing Hao	24.98	–
Huang Lian	25.26	–
Cang Er Zi	26.63	–
Wu Wei Zi	27.14	–
Qiang Huo	27.88	–
Xi Xian Cao	28.40	–
Du Zhong	29.00	–
Zi Hua Di Ding	30.60	–
Ze Lan	30.63	–
Jin Yin Hua	30.72	–
Mao Dong Qing	30.85	–
Jiao Gu Lan	31.30	–
Pu Gong Ying	31.62	–
Nü Zhen Zi	32.78	–
Hou Po	32.79	–
Xin Yi	33.14	–
Hong Hua	33.35	–
Jiang Huang	33.44	–
Sha Ren	33.75	–
Tu Fu Ling	33.98	–
Zhi Shi	34.08	–
Yin Chen Hao	34.92	–
Hong Jing Tian	35.54	–
Bai Jiang Cao	35.67	–
Niu Bang Zi	35.76	–
Yi Mu Cao	36.48	–
Gan Cao	37.19	–
Jing Jie	37.88	–
Dan Dou Chi	37.96	–
Lian Qiao	38.55	–
Yin Yang Huo	39.02	–
Zhi Ke	40.04	–
Fu Zi	40.29	–
Bo He	40.43	–
Guang Huo Xiang	40.90	–
Dan Shen	42.21	–
Shen Qu	42.82	–
Ma Huang	43.39	–
Wang Bu Liu Xing	44.56	–
Shan Yu Rou	45.39	–
Ye Jiao Teng	45.40	–
Chai Hu	45.53	–
He Huan Pi	45.65	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gua Lou	46.62	–
Ce Bai Ye	46.85	–
Hua Shi	46.88	–
Xie Bai	47.67	–
Hu Zhang	47.78	–
Zi Su Zi	48.67	–
Zhe Bei Mu	48.75	–
Jie Geng	49.20	–
Ku Shen	49.26	–
Yu Xing Cao	50.33	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wu Mei* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62477	62477	0.00	24.34
62478	62478	0.00	24.29
62715	62715	0.00	16.45
62716	62716	0.00	14.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wu Wei Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60001-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wu Wei Zi; Schisandrae chinensis fructus

Special notes

When selecting the *Wu Wei Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wu Wei Zi	3	0	1

Second-stage model

For differentiation of the substance/substance group *Wu Wei Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wu Wei Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wu Wei Zi	G220HS048SK1	62551	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048SK1	62552	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048SN1	62705	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048SN1	62706	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048TK1	62783	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048TK1	62784	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Wu Wei Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Wu Wei Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wu Wei Zi	G220HS048SK1	62551†	20
PhytoComm	Wu Wei Zi	G220HS048SK1	62552†	20
PhytoComm	Wu Wei Zi	G220HS048SN1	62705†	20
PhytoComm	Wu Wei Zi	G220HS048SN1	62706†	20
PhytoComm	Wu Wei Zi	G220HS048TK1	62783†	20
PhytoComm	Wu Wei Zi	G220HS048TK1	62784†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Wu Wei Zi*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Caelo	Wu Wei Zi	G220H0401421	2
Phytocomm	Wu Wei Zi	G220H0401421	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wu Wei Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wu Wei Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	0	0	3	854

The substance/substance group *Wu Wei Zi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wu Wei Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	10.76	–
Jiao Gu Lan	12.21	–
Shan Yao	13.52	–
Zhi Ke	15.45	–
Ku Shen	15.53	–
Suan Zao Ren	16.05	–
Lian Qiao	18.08	–
Cang Zhu	18.94	–
Shan Yu Rou	20.65	–
Xie Bai	21.30	–
Yan Hu Suo	22.34	–
San Qi	23.66	–
Chuan Xiong	24.20	–
Dan Shen	25.93	–
Zhe Bei Mu	26.34	–
Ling Zhi	26.90	–
Wu Mei	26.95	–
Hou Po	28.83	–
Qing Pi	29.08	–
Fu Zi	29.29	–
Mao Dong Qing	29.63	–
Zi Hua Di Ding	29.63	–
Shen Qu	30.32	–
Bai Zi Ren	30.59	–
Qing Hao	30.65	–
Gou Qi Zi	30.86	–
Cang Er Zi	30.93	–
Ji Li	30.99	–
Ze Xie	31.18	–
Pi Pa Ye	31.26	–
He Huan Pi	31.33	–
Jie Geng	31.81	–
Di Gu Pi	32.04	–
Sang Ye	32.47	–
Dan Dou Chi	32.63	–
Zhi Gan Cao	32.90	–
Tai Zi Shen	33.11	–
Mang Xiao	33.14	–
Guang Huo Xiang	33.15	–
Du Zhong	33.17	–
Huang Bai	33.38	–
Gua Lou	33.40	–
Ye Jiao Teng	33.42	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Zhu	34.12	–
Fu Pen Zi	34.37	–
Chen Pi	34.41	–
Zhu Ru	34.73	–
Tian Hua Fen	35.80	–
Hong Hua	35.87	–
She Gan	36.24	–
Dang Gui	36.51	–
Tao Ren	36.66	–
Qiang Huo	36.74	–
Ren Dong Teng	37.01	–
Yin Yang Huo	37.20	–
Chai Hu	37.23	–
Lai Fu Zi	37.75	–
Huo Ma Ren	37.84	–
Gan Cao	37.87	–
Huang Lian	38.12	–
Ce Bai Ye	38.42	–
Che Qian Zi	38.46	–
Sheng Jiang	38.48	–
Chi Shao (Yao)	38.55	–
Bai Shao Yao	38.91	–
Mu Gua	39.53	–
Yi Mu Cao	39.54	–
Ren Shen	39.65	–
Lian Zi	39.94	–
Mu Zei	40.06	–
Lu Gen	40.33	–
Chuang Mu Xiang	40.73	–
Yuan Zhi	41.26	–
Ban Lan Gen	41.38	–
Gu Sui Bu	41.57	–
Yi Yi Ren	41.73	–
Huang Qin	41.94	–
Bo He	43.39	–
Ji Xue Teng	44.05	–
Bai Xian Pi	44.33	–
Zhu Ling	45.23	–
Jiang Huang	45.35	–
Fu Ling	45.46	–
Xi Xian Cao	45.79	–
Rou Gui	46.67	–
Jing Jie	47.05	–
Sha Ren	47.16	–
Dang Gui Wei	47.31	–
Nü Zhen Zi	47.42	–
Zhi Mu	47.61	–
Ma Huang	48.74	–
Gui Zhi	48.86	–
Sang Zhi	49.05	–
Gou Teng	49.15	–
Yu Jin	49.79	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wu Wei Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62551	62551	0.00	11.85
62552	62552	0.00	10.76
62705	62705	0.00	27.69
62706	62706	0.00	26.78
62783	62783	0.00	30.32
62784	62784	0.00	30.57

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wu Yao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60040-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wu Yao; Linderae radix

Special notes

When selecting the *Wu Yao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wu Yao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Wu Yao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wu Yao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wu Yao	G140H1025921	62741	40	from supplier
PhytoComm	Wu Yao	G140H1025921	62742	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Wu Yao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Wu Yao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wu Yao	G140H1025921	62741†	20
PhytoComm	Wu Yao	G140H1025921	62742†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 6 *Apo-Ident* customers from 3 batches from the substance/substance group *Wu Yao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Wu Yao	g140h1025221	1
PhytoComm	Wu Yao	G140H1025221	2
Phytocomm	Wu Yao	G140H1025521	2
PhytoComm	Wu Yao	G140H1025521	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wu Yao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wu Yao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	6	851

The substance/substance group *Wu Yao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wu Yao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
(Shi) Chang Pu	11.83	–
Chen Pi	13.42	–
He Huan Pi	15.54	–
Ji Li	17.10	–
Jiang Huang	17.33	–
Sha Ren	19.00	–
Shan Yu Rou	20.20	–
Gu Sui Bu	20.71	–
Mu Dan Pi	20.87	–
Chuan Lian Zi	22.10	–
(Bai) Dou Kou	22.29	–
Ju Hua	25.21	–
Yin Chen Hao	25.48	–
Dang Gui	27.29	–
Mang Xiao	27.42	–
Yan Hu Suo	27.61	–
Hong Jing Tian	27.77	–
Sang Ji Shend	28.25	–
He Shou Wu	29.17	–
Yi Mu Cao	29.84	–
Xiang Fu	30.25	–
Shan Yao	30.57	–
Xiao Hui Xiang	30.60	–
Di Gu Pi	30.73	–
Ba Ji Tian	31.25	–
Sang Bai Pi	31.50	–
Wu Zhu Yu	31.83	–
E Zhu	32.98	–
Tu Si Zi	33.35	–
Fang Feng	34.07	–
Chuan Niu Xi	34.09	–
Niu Bang Zi	36.59	–
Xu Duan	37.13	–
Huang Qi	37.29	–
Bu Gu Zhi	38.76	–
Bai Hua She She Cao	38.79	–
(Huai) Niu Xi	39.06	–
Gua Lou	39.31	–
Sang Zhi	39.35	–
Zi Su Zi	39.53	–
Dan Zhu Ye	39.68	–
Ku Shen	40.12	–
Zhi Gan Cao	40.53	–
Zhi Ke	40.78	–
Gan Jiang	40.98	–
Ge Gen	41.19	–
Tian Hua Fen	41.26	–
Bai Zhi	41.92	–
Xin Yi	41.97	–
Du Huo	42.18	–
Huang Lian	42.41	–
Hu Zhang	43.66	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lai Fu Zi	44.17	–
Cang Er Zi	45.26	–
Suan Zao Ren	45.96	–
Wang Bu Liu Xing	46.78	–
Ze Lan	47.53	–
Nü Zhen Zi	47.79	–
Da Zao	48.19	–
Jie Geng	49.82	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wu Yao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62741	62741	0.00	11.84
62742	62742	0.00	11.83

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Wu Zhu Yu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002260-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Wu Zhu Yu; Evodiae rutecarpae fructus

Special notes

When selecting the *Wu Zhu Yu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.
2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Wu Zhu Yu	1	0	2

Second-stage model

For differentiation of the substance/substance group *Wu Zhu Yu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Wu Zhu Yu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wu Zhu Yu	G103H0766921	62781	40	from supplier
PhytoComm	Wu Zhu Yu	G103H0766921	62782	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Wu Zhu Yu*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Wu Zhu Yu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Wu Zhu Yu	G103H0766921	62781†	20
PhytoComm	Wu Zhu Yu	G103H0766921	62782†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Wu Zhu Yu*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Wu Zhu Yu	G103H0766022	1
PhytoComm	Wu Zhu Yu	G103H0766321	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Wu Zhu Yu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Wu Zhu Yu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	2	1	1	853

The substance/substance group *Wu Zhu Yu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.4419 % (> 98.8569 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Wu Zhu Yu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
(Shi) Chang Pu	19.45	–
Xiao Hui Xiang	19.71	–
Yi Mu Cao	21.86	–
Nü Zhen Zi	22.63	–
Ju Hua	23.63	–
Gu Sui Bu	24.12	–
Sang Ji Shend	24.95	–
Sang Bai Pi	25.97	–
Sha Ren	26.44	–
Mang Xiao	26.69	–
Shan Yu Rou	27.43	–
Ku Shen	27.49	–
Dan Zhu Ye	27.50	–
Fang Feng	29.07	–
Xu Duan	30.45	–
Yin Chen Hao	30.59	–
Xiang Fu	30.64	–
Hong Jing Tian	31.12	–
Chen Pi	31.19	–
Lian Qiao	32.08	–
He Huan Pi	34.76	–
Wu Yao	36.86	–
Cang Er Zi	37.27	–
Bai Hua She She Cao	38.24	–
Dang Gui	39.13	–
Mu Dan Pi	40.31	–
Jing Jie	40.31	–
Di Gu Pi	40.66	–
Sang Ye	41.85	–
He Shou Wu	43.08	–
Jiang Huang	43.48	–
Gou Teng	43.48	–
Gua Lou	44.16	–
Rou Cong Rong	44.63	–
Shan Yao	46.04	–
Bai Zhi	46.12	–
(Bai) Dou Kou	46.19	–
Ge Gen	46.78	–
Yan Hu Suo	46.84	–
Niu Bang Zi	46.96	–
Xuan Fu Hua	47.46	–
Pu Gong Ying	48.01	–
Tu Si Zi	49.25	–
Han Lian Cao	49.77	–
Xin Yi	50.37	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wu Zhu Yu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62781	62781	0.00	19.45
62782	62782	0.00	19.71

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xi Xian Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60103-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xi Xian Cao; Siegesbeckiae herba

Special notes

When selecting the *Xi Xian Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xi Xian Cao	2	0	2

Second-stage model

For differentiation of the substance/substance group *Xi Xian Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xi Xian Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xi Xian Cao	G229HS309SK1	62555	40	from supplier
PhytoComm	Xi Xian Cao	G229HS309SK1	62556	40	from supplier
PhytoComm	Xi Xian Cao	G229HS309SQ1	62791	40	from supplier
PhytoComm	Xi Xian Cao	G229HS309SQ1	62792	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Xi Xian Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Xi Xian Cao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xi Xian Cao	G229HS309SK1	62555†	20
PhytoComm	Xi Xian Cao	G229HS309SK1	62556†	20
PhytoComm	Xi Xian Cao	G229HS309SQ1	62791†	20
PhytoComm	Xi Xian Cao	G229HS309SQ1	62792†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Xi Xian Cao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xi Xian Cao	G229H1431321	1
PhytoComm	Xi Xian Cao	G229H1431321	1
PhytoComm	Xi Xian Cao	G229H1431521	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xi Xian Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xi Xian Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	79	1	12 397
Type C	0	0	3	854

The substance/substance group *Xi Xian Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xi Xian Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ze Lan	4.83	–
Tu Fu Ling	7.64	–
Yin Chen Hao	7.78	–
Xin Yi	9.31	–
Huang Lian	13.25	–
Yu Xing Cao	14.13	–
Pu Gong Ying	14.57	–
Nü Zhen Zi	15.88	–
Bai Hua She She Cao	16.56	–
Shan Yao	18.39	–
Du Zhong	19.07	–
Sang Ye	19.36	–
Yu Jin	20.17	–
Hu Zhang	21.39	–
Jiang Huang	21.72	–
Sha Ren	22.29	–
Zhi Ke	23.34	–
Qiang Huo	23.42	–
Zhi Shi	23.84	–
Bai Jiang Cao	24.29	–
Fu Zi	24.47	–
Han Lian Cao	25.85	–
Jiao Gu Lan	25.87	–
Dan Shen	27.89	–
Wu Mei	28.38	–
Hou Po	28.68	–
Mao Dong Qing	28.79	–
Bo He	28.91	–
Mang Xiao	29.31	–
Che Qian Zi	30.78	–
Bai Xian Pi	30.98	–
Cang Er Zi	31.99	–
Qing Pi	32.24	–
Ling Zhi	32.39	–
Yin Yang Huo	32.53	–
Chai Hu	32.89	–
Xiang Fu	32.92	–
Zi Hua Di Ding	33.72	–
Wang Bu Liu Xing	33.88	–
Ye Jiao Teng	33.90	–
Yan Hu Suo	33.99	–
Huang Bai	34.34	–
Dan Dou Chi	34.84	–
Hong Jing Tian	35.23	–
He Huan Pi	35.75	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Di Gu Pi	36.21	–
Ce Bai Ye	36.85	–
Wu Wei Zi	37.32	–
Jin Yin Hua	37.63	–
Yi Mu Cao	37.83	–
E Zhu	39.56	–
Fu Pen Zi	40.07	–
Zi Su Zi	40.20	–
Xia Ku Cao	40.38	–
Qing Hao	40.91	–
Chuang Mu Xiang	40.96	–
Lian Zi	41.14	–
Pi Pa Ye	41.14	–
Ma Huang	41.78	–
Ji Li	42.14	–
Jing Jie	43.53	–
(Shi) Chang Pu	45.40	–
Lian Qiao	45.85	–
Ren Dong Teng	46.06	–
Chuan Xiong	46.27	–
Suan Zao Ren	46.41	–
Gua Lou	46.41	–
Sang Ji Shend	46.47	–
Niu Bang Zi	47.17	–
Bai Shao Yao	47.52	–
(Bai) Dou Kou	47.58	–
Ji Xue Teng	48.67	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xi Xian Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62555	62555	0.00	7.78
62556	62556	0.00	7.65
62791	62791	0.00	4.83
62792	62792	0.00	5.75

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xia Ku Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60080-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xia Ku Cao; Prunellae vulgaris spica

Special notes

When selecting the *Xia Ku Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xia Ku Cao	1	0	3

Second-stage model

For differentiation of the substance/substance group *Xia Ku Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xia Ku Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xia Ku Cao	G202HS232RM1	62283	40	from supplier
PhytoComm	Xia Ku Cao	G202HS232RM1	62284	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xia Ku Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.

- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xia Ku Cao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xia Ku Cao	G202HS232RM1	62283†	20
PhytoComm	Xia Ku Cao	G202HS232RM1	62284†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Xia Ku Cao*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xia Ku Cao	g202h1064121	1
Phytocomm	Xia Ku Cao	G202H1064221	2
PhytoComm	Xia Ku Cao	G202H1064221	2
Phytocomm	Xia Ku Cao	G202H1064421	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xia Ku Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xia Ku Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	6	851

The substance/substance group *Xia Ku Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xia Ku Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Zhi Shi	7.05	–
Xin Yi	16.27	–
Hu Zhang	19.96	–
Bai Hua She She Cao	22.28	–
Han Lian Cao	23.46	–
Mang Xiao	25.93	–
Sang Ye	26.43	–
Yi Mu Cao	27.03	–
Xi Xian Cao	36.07	–
Yu Xing Cao	38.30	–
Sang Ji Shend	38.33	–
He Shou Wu	39.38	–
Huang Lian	39.81	–
Gu Sui Bu	40.87	–
He Huan Pi	41.05	–
Sha Ren	41.82	–
Xuan Fu Hua	41.90	–
Gou Teng	42.21	–
Jiang Huang	42.26	–
Xuan Shen	43.78	–
Nü Zhen Zi	44.15	–
Sang Bai Pi	44.39	–
Bu Gu Zhi	44.97	–
Jing Jie	47.33	–
Dan Zhu Ye	47.79	–
Pu Gong Ying	47.82	–
Rou Cong Rong	48.99	–
Ze Lan	49.06	–
Shan Yao	49.96	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xia Ku Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62283	62283	0.00	7.10
62284	62284	0.00	7.05

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at

least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xian Mao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60316-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xian Mao; Curculiginis orchioidis rhizoma

Special notes

When selecting the *Xian Mao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xian Mao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Xian Mao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xian Mao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xian Mao	G304H0549021	62985	40	from supplier
PhytoComm	Xian Mao	G304H0549021	62986	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xian Mao*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xian Mao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xian Mao	G304H0549021	62985†	20
PhytoComm	Xian Mao	G304H0549021	62986†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Xian Mao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xian Mao	G304H0549321	2
Phytocomm	Xian Mao	G304H0549521	2

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xian Mao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xian Mao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	4	853

The substance/substance group *Xian Mao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xian Mao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Wu Jia Pi	8.05	–
Yu Xing Cao	9.97	–
(Sheng) Di Huang	15.11	–
Dan Zhu Ye	17.24	–
Mang Xiao	22.42	–
Jing Jie	26.43	–
Ding Xiang	28.89	–
Sang Ji Shend	29.92	–
Sang Ye	30.01	–
Shu Di (Huang)	34.60	–
Bu Gu Zhi	37.39	–
Ge Gen	38.60	–
Du Zhong	43.38	–
Hua Shi	46.64	–
Jin Qian Cao	48.39	–
Yin Chen Hao	49.89	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xian Mao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62985	62985	0.00	8.05
62986	62986	0.00	8.96

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xiang Fu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60185-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xiang Fu; Cyperi rotundi rhizoma

Special notes

When selecting the *Xiang Fu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xiang Fu	2	0	3

Second-stage model

For differentiation of the substance/substance group *Xiang Fu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xiang Fu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xiang Fu	G088H0930921	62777	40	from supplier
PhytoComm	Xiang Fu	G088H0930921	62778	40	from supplier
PhytoComm	Xiang Fu	G088H0930922	62937	40	from supplier
PhytoComm	Xiang Fu	G088H0930922	62938	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Xiang Fu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Xiang Fu*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xiang Fu	G088H0930921	62777†	20
PhytoComm	Xiang Fu	G088H0930921	62778†	20
PhytoComm	Xiang Fu	G088H0930922	62937†	20
PhytoComm	Xiang Fu	G088H0930922	62938†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 8 spectra from 5 *Apo-Ident* customers from 3 batches from the substance/substance group *Xiang Fu*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xiang Fu	g088h0930123	1
PhytoComm	Xiang Fu	G088H0930321	1
Phytocomm	Xiang Fu	G088H0930521	5
PhytoComm	Xiang Fu	G088H0930521	1

- 849 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xiang Fu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xiang Fu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	1	0	8	848

The substance/substance group *Xiang Fu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.9483 % (> 99.3610 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xiang Fu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yi Mu Cao	14.47	–
Gua Lou	16.40	–
Zi Su Zi	16.98	–
Niu Bang Zi	18.50	–
Wu Yao	18.77	–
(Shi) Chang Pu	19.34	–
Cang Er Zi	19.94	–
Di Gu Pi	20.56	–
Shan Yao	21.31	–
Jiang Huang	21.40	–
Xiao Hui Xiang	21.91	–
Yan Hu Suo	22.16	–
Ku Shen	23.11	–
Gu Sui Bu	23.67	–
Sha Ren	23.80	–
Zhi Gan Cao	23.97	–
Tian Hua Fen	24.34	–
Lian Qiao	24.66	–
Ju Hua	24.66	–
(Huai) Niu Xi	24.69	–
Xin Yi	25.15	–
Bai Hua She She Cao	25.15	–
Pu Gong Ying	25.80	–
Chuan Lian Zi	25.86	–
Chen Pi	26.05	–
Bai Zhi	26.13	–
Huang Qi	26.56	–
Chuan Niu Xi	26.75	–
(Bai) Dou Kou	26.88	–
Chi Shao (Yao)	27.01	–
Zhi Ke	27.31	–
Shan Yu Rou	27.48	–
Dang Gui	27.54	–
Sang Ye	27.94	–
Mang Xiao	28.00	–
Xu Duan	28.28	–
Jie Geng	28.35	–
Wu Wei Zi	28.78	–
Ze Lan	29.28	–
Ji Li	29.56	–
Yuan Zhi	29.73	–
Bo He	30.99	–
Sang Zhi	31.42	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xi Xian Cao	32.37	–
Wu Zhu Yu	32.41	–
Hu Zhang	32.45	–
Gan Cao	32.67	–
Jin Yin Hua	32.89	–
He Huan Pi	33.45	–
Gan Jiang	34.13	–
Huang Lian	34.23	–
Tu Fu Ling	34.61	–
Mao Dong Qing	34.90	–
Chuan Mu Tong	35.19	–
Zi Hua Di Ding	35.30	–
E Zhu	36.19	–
Du Zhong	36.70	–
Bai Zhu	36.77	–
Nü Zhen Zi	37.60	–
Qiang Huo	38.04	–
Mu Gua	38.37	–
Dong Gua Zi	38.38	–
Mu Dan Pi	38.91	–
Fang Feng	39.47	–
Sang Ji Shend	39.83	–
Hou Po	39.93	–
Mi Huan Jun	40.10	–
Du Huo	40.73	–
Han Lian Cao	40.73	–
Fu Zi	40.76	–
Qing Pi	40.78	–
Bing Lang	41.15	–
Lian Zi	41.75	–
Yu Xing Cao	41.91	–
Yin Chen Hao	42.02	–
Ce Bai Ye	42.09	–
Huang Bai	42.13	–
Jiao Gu Lan	42.31	–
Dan Shen	42.43	–
Bai Jiang Cao	42.56	–
Yu Jin	42.63	–
Zhi Shi	42.73	–
Wang Bu Liu Xing	42.80	–
Jing Jie	42.92	–
Chuang Mu Xiang	43.05	–
Cang Zhu	43.16	–
Suan Zao Ren	43.32	–
Sha Shen (Bei)	43.55	–
Chai Hu	43.85	–
Chuan Xiong	43.95	–
Dan Zhu Ye	44.18	–
He Shou Wu	44.44	–
Dan Dou Chi	44.54	–
Gou Qi Zi	44.77	–
Hong Jing Tian	45.70	–
Che Qian Zi	45.75	–
Bai Xian Pi	45.75	–
Qin Jiao	46.17	–
Qing Hao	46.21	–
Ye Jiao Teng	46.36	–
Shen Qu	46.97	–
Ba Ji Tian	47.81	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Mei	47.85	–
Xia Ku Cao	48.03	–
Ge Gen	48.03	–
Lai Fu Zi	48.21	–
Bai He	48.25	–
Zhe Bei Mu	48.89	–
Ren Dong Teng	49.65	–
Sang Bai Pi	49.80	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xiang Fu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62777	62777	0.00	17.19
62778	62778	0.00	16.98
62937	62937	0.00	14.51
62938	62938	0.00	14.47

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xiao Hui Xiang**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60209-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xiao Hui Xiang; Foeniculi vulgaris fructus

Special notes

When selecting the *Xiao Hui Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xiao Hui Xiang	1	0	2

Second-stage model

For differentiation of the substance/substance group *Xiao Hui Xiang* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xiao Hui Xiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xiao Hui Xiang	G104H0339021	63003	40	from supplier
PhytoComm	Xiao Hui Xiang	G104H0339021	63004	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xiao Hui Xiang*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xiao Hui Xiang*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xiao Hui Xiang	G104H0339021	63003†	20
PhytoComm	Xiao Hui Xiang	G104H0339021	63004†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 4 *Apo-Ident* customers from 2 batches from the substance/substance group *Xiao Hui Xiang*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xiao Hui Xiang	g104h0339321	1
Phytocomm	Xiao Hui Xiang	G104H0339522	3

- 853 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xiao Hui Xiang* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xiao Hui Xiang* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	4	853

The substance/substance group *Xiao Hui Xiang* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xiao Hui Xiang* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ku Shen	18.30	–
Yi Mu Cao	18.44	–
Xu Duan	21.86	–
Xiang Fu	23.58	–
Wu Zhu Yu	25.13	–
Nü Zhen Zi	25.88	–
Lian Qiao	27.09	–
Cang Er Zi	27.15	–
Sha Ren	27.41	–
Gua Lou	28.07	–
(Shi) Chang Pu	28.64	–
Niu Bang Zi	28.65	–
Mang Xiao	29.22	–
Yan Hu Suo	30.60	–
Shan Yu Rou	32.20	–
(Bai) Dou Kou	32.55	–
Jiang Huang	33.00	–
Ju Hua	33.86	–
Gu Sui Bu	35.52	–
Chen Pi	36.79	–
Shan Yao	36.88	–
Sang Ji Shend	38.16	–
Fang Feng	38.58	–
Ze Lan	38.84	–
He Huan Pi	39.05	–
Zi Su Zi	39.31	–
Wu Yao	40.38	–
Yin Chen Hao	40.88	–
Bai Hua She She Cao	41.14	–
Shen Qu	41.21	–
Suan Zao Ren	42.47	–
Di Gu Pi	42.60	–
Sang Ye	42.67	–
Sang Bai Pi	43.38	–
(Huai) Niu Xi	43.59	–
Dan Zhu Ye	43.61	–
Pu Gong Ying	43.87	–
Rou Cong Rong	44.52	–
Hong Jing Tian	44.65	–
Han Lian Cao	45.98	–
Dang Gui	47.30	–
Jing Jie	47.59	–
Bai Zhi	47.78	–
Wang Bu Liu Xing	47.85	–
Qiang Huo	49.22	–
Chuan Niu Xi	49.24	–
Gan Cao	49.83	–
Xi Xian Cao	50.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xiao Hui Xiang* is separated from critical neighbours in a second-stage model,

all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
63003	63003	0.00	18.48
63004	63004	0.00	18.30

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xie Bai**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60362-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xie Bai; Allii bulbus

Special notes

When selecting the *Xie Bai* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xie Bai	1	0	3

Second-stage model

For differentiation of the substance/substance group *Xie Bai* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xie Bai*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xie Bai	G504HS334TK1	62819	40	from supplier
PhytoComm	Xie Bai	G504HS334TK1	62820	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xie Bai*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xie Bai*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xie Bai	G504HS334TK1	62819†	20
PhytoComm	Xie Bai	G504HS334TK1	62820†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 2 *Apo-Ident* customers from 3 batches from the substance/substance group *Xie Bai*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Xie Bai	G504H1830321	1
Phytocomm	Xie Bai	G504H1830421	1
Phytocomm	Xie Bai	G504H1830422	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xie Bai* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xie Bai* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	3	854

The substance/substance group *Xie Bai* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xie Bai* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Cang Zhu	11.07	–
Yuan Zhi	12.86	–
Chen Pi	13.35	–
Chuang Mu Xiang	14.06	–
Shan Yao	15.12	–
Yu Zhu	15.64	–
Suan Zao Ren	16.27	–
Chuan Xiong	18.15	–
Lian Qiao	18.25	–
San Qi	18.31	–
Gou Qi Zi	18.59	–
Bai Shao Yao	18.85	–
Mu Gua	18.90	–
Jiao Gu Lan	19.04	–
Lian Zi	19.09	–
Jie Geng	19.20	–
Pi Pa Ye	19.29	–
Bai Zi Ren	19.49	–
Shan Yu Rou	19.62	–
Dang Gui Wei	19.75	–
Ku Shen	20.25	–
Ji Li	20.73	–
Zhe Bei Mu	21.73	–
Tian Hua Fen	21.91	–
Zhi Mu	22.20	–
Tai Zi Shen	22.28	–
Mao Dong Qing	23.43	–
Ren Shen	23.52	–
Ling Zhi	23.90	–
Shen Qu	24.54	–
Dang Gui	24.64	–
Jin Yin Hua	25.03	–
Dan Dou Chi	25.26	–
Di Gu Pi	25.27	–
Zhi Ke	25.67	–
Ze Xie	25.82	–
Gua Lou	25.86	–
Dan Shen	26.27	–
Guang Huo Xiang	27.65	–
Wu Wei Zi	27.81	–
Chai Hu	27.89	–
Hong Jing Tian	27.97	–
Mu Zei	28.48	–
Long Yan Rou	28.54	–
Zhu Ru	28.87	–
Lai Fu Zi	29.30	–
Huo Ma Ren	29.62	–
Sheng Jiang	29.72	–
Ban Lan Gen	29.80	–
Ren Dong Teng	29.88	–
Hou Po	29.96	–
He Huan Pi	30.44	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gu Sui Bu	31.65	–
Fu Ling	31.82	–
Lu Gen	32.22	–
Ci Wu Jia	32.81	–
Mang Xiao	33.16	–
Tu Fu Ling	33.72	–
She Gan	33.74	–
Ye Jiao Teng	33.84	–
Fu Zi	34.52	–
Gan Cao	34.62	–
Bai Xian Pi	35.05	–
Gui Zhi	35.28	–
Gou Teng	36.12	–
Cang Er Zi	36.24	–
Hong Hua	36.28	–
Fu Pen Zi	37.05	–
Mai Men Dong	37.11	–
Zi Hua Di Ding	37.13	–
Tao Ren	37.19	–
Rou Gui	37.40	–
Zhi Gan Cao	38.10	–
Yan Hu Suo	38.41	–
Ji Xue Teng	39.28	–
Fo Shou	40.57	–
Qing Pi	40.81	–
Fu Xiao Mai	41.23	–
Yin Yang Huo	41.55	–
Bo He	41.79	–
Huang Qin	42.03	–
Ma Huang	42.04	–
Che Qian Zi	42.05	–
Qiang Huo	42.38	–
Ce Bai Ye	43.06	–
Yu Jin	43.19	–
Yi Yi Ren	43.31	–
Sang Zhi	43.39	–
Chi Shao (Yao)	43.50	–
Ma Huang Gen	46.63	–
Sha Ren	46.98	–
Zhu Ling	47.11	–
Huang Bai	47.27	–
Du Zhong	47.59	–
Chuan Lian Zi	48.89	–
(Fen) Bi Xie	49.23	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xie Bai* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62819	62819	0.00	11.07
62820	62820	0.00	11.83

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xin Yi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60231-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xin Yi; Magnoliae flos

Special notes

When selecting the *Xin Yi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xin Yi	3	0	2

Second-stage model

For differentiation of the substance/substance group *Xin Yi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xin Yi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xin Yi	G153HS139SH1	62603	40	from supplier
PhytoComm	Xin Yi	G153HS139SH1	62604	40	from supplier
PhytoComm	Xin Yi	G153HS139SP1	62725	40	from supplier
PhytoComm	Xin Yi	G153HS139SP1	62726	40	from supplier
PhytoComm	Xin Yi	G153HS139TK1	62885	40	from supplier
PhytoComm	Xin Yi	G153HS139TK1	62886	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Xin Yi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Xin Yi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xin Yi	G153HS139SH1	62603†	20
PhytoComm	Xin Yi	G153HS139SH1	62604†	20
PhytoComm	Xin Yi	G153HS139SP1	62725†	20
PhytoComm	Xin Yi	G153HS139SP1	62726†	20
PhytoComm	Xin Yi	G153HS139TK1	62885†	20
PhytoComm	Xin Yi	G153HS139TK1	62886†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Xin Yi*.

- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Xin Yi	G153H0758321	1
Phytocomm	Xin Yi	G153H0758521	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xin Yi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xin Yi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	119	1	12 357
Type C	0	0	2	855

The substance/substance group *Xin Yi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	99.1667 % (> 96.6667 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xin Yi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Hua She She Cao	5.31	–
Zhi Shi	6.32	–
Xi Xian Cao	6.67	–
Pu Gong Ying	7.38	–
Xia Ku Cao	9.64	–
Ze Lan	10.39	–
Han Lian Cao	10.64	–
Tu Fu Ling	11.11	–
Yu Xing Cao	11.94	–
Yin Chen Hao	13.54	–
Nü Zhen Zi	14.08	–
Hu Zhang	14.80	–
Sang Ye	14.85	–
Shan Yao	17.43	–
Huang Lian	18.24	–
Jiang Huang	19.43	–
Yi Mu Cao	20.79	–
Du Zhong	24.79	–
Rou Cong Rong	25.06	–
Yu Jin	26.55	–
Sha Ren	26.65	–
Mang Xiao	26.65	–
Qiang Huo	26.92	–
Fu Zi	28.81	–
He Huan Pi	29.66	–
Xiang Fu	31.13	–
Bai Jiang Cao	31.82	–
Dan Shen	32.25	–
Wu Mei	32.33	–
Gou Teng	33.07	–
Zhi Ke	33.12	–
Hou Po	33.57	–
Gu Sui Bu	33.68	–
Cang Er Zi	34.25	–
Sang Ji Shend	34.62	–
Bo He	34.87	–
Yan Hu Suo	35.13	–
Che Qian Zi	36.33	–
Mao Dong Qing	36.72	–
Qing Pi	36.87	–
Wang Bu Liu Xing	37.10	–
Jiao Gu Lan	37.97	–
Hong Jing Tian	38.15	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yin Yang Huo	38.97	–
Zi Hua Di Ding	39.02	–
Xuan Shen	39.23	–
Dan Dou Chi	39.25	–
Huang Bai	39.86	–
Bai Xian Pi	40.03	–
Chai Hu	40.27	–
Wu Wei Zi	41.58	–
Ye Jiao Teng	41.74	–
Zi Su Zi	42.14	–
Jing Jie	42.48	–
Di Gu Pi	42.53	–
Ling Zhi	42.58	–
Gua Lou	42.59	–
E Zhu	43.70	–
Jin Yin Hua	43.91	–
Ce Bai Ye	44.62	–
Ji Li	45.29	–
Fu Pen Zi	46.13	–
Niu Bang Zi	46.24	–
Qing Hao	46.48	–
Dan Zhu Ye	48.09	–
Chuang Mu Xiang	48.23	–
Ju Hua	48.44	–
Bu Gu Zhi	48.51	–
Sang Bai Pi	48.69	–
Wu Zhu Yu	48.81	–
Ma Huang	49.05	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xin Yi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62603	62603	0.00	6.71
62604	62604	0.00	6.67
62725	62725	0.00	14.29
62726	62726	0.00	14.93
62885	62885	0.00	5.90
62886	62886	0.00	5.31

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances,

thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xu Duan**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60245-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xu Duan; Dipsaci radix

Special notes

When selecting the *Xu Duan* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xu Duan	1	0	1

Second-stage model

For differentiation of the substance/substance group *Xu Duan* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xu Duan*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xu Duan	G092H2106922	62911	40	from supplier
PhytoComm	Xu Duan	G092H2106922	62912	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xu Duan*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xu Duan*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xu Duan	G092H2106922	62911†	20
PhytoComm	Xu Duan	G092H2106922	62912†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Xu Duan*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xu Duan	G092H2106422	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xu Duan* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xu Duan* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Xu Duan* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xu Duan* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ku Shen	12.87	–
Yi Mu Cao	16.47	–
Gua Lou	17.53	–
Xiang Fu	17.72	–
Xiao Hui Xiang	23.48	–
Chen Pi	24.53	–
Cang Er Zi	25.12	–
(Huai) Niu Xi	25.18	–
Shan Yu Rou	25.64	–
Lian Qiao	26.51	–
Chuan Niu Xi	27.29	–
Yan Hu Suo	28.48	–
(Shi) Chang Pu	28.60	–
Mang Xiao	28.80	–
Wu Zhu Yu	29.11	–
Niu Bang Zi	30.98	–
Wu Yao	31.52	–
Huang Qi	32.19	–
Fang Feng	32.51	–
Shan Yao	32.87	–
Zi Su Zi	34.18	–
Gan Cao	34.23	–
Chuan Lian Zi	34.46	–
Gu Sui Bu	34.87	–
Di Gu Pi	34.94	–
Bai Zhi	35.34	–
Tian Hua Fen	36.36	–
Ju Hua	36.71	–
Pu Gong Ying	37.35	–
Sha Ren	37.63	–
Chi Shao (Yao)	38.30	–
Dang Gui	38.33	–
Yuan Zhi	38.35	–
Jie Geng	38.74	–
Jiang Huang	39.22	–
(Bai) Dou Kou	39.57	–
Zhi Ke	39.86	–
Wu Wei Zi	40.01	–
Zhi Gan Cao	40.02	–
Shen Qu	40.55	–
He Huan Pi	40.95	–
Sang Zhi	41.34	–
Sang Ye	42.76	–
Jin Yin Hua	42.82	–
Xin Yi	43.69	–
Nü Zhen Zi	43.95	–
Ze Lan	43.99	–
Bai Hua She She Cao	44.51	–
Dong Gua Zi	44.82	–
Ba Ji Tian	46.69	–
Sang Ji Shend	47.30	–
Bai Zhu	47.48	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Bai Pi	48.44	–
Hong Jing Tian	48.48	–
Bo He	48.60	–
Ji Li	48.79	–
Rou Cong Rong	49.09	–
Huang Lian	49.61	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xu Duan* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62911	62911	0.00	12.96
62912	62912	0.00	12.87

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xuan Fu Hua**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60223-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xuan Fu Hua; Inulae flos

Special notes

When selecting the *Xuan Fu Hua* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xuan Fu Hua	1	0	1

Second-stage model

For differentiation of the substance/substance group *Xuan Fu Hua* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xuan Fu Hua*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xuan Fu Hua	G126H1168921	62901	40	from supplier
PhytoComm	Xuan Fu Hua	G126H1168921	62902	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Xuan Fu Hua*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Xuan Fu Hua*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xuan Fu Hua	G126H1168921	62901†	20
PhytoComm	Xuan Fu Hua	G126H1168921	62902†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Xuan Fu Hua*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Xuan Fu Hua	G126H1168521	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xuan Fu Hua* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xuan Fu Hua* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	1	856

The substance/substance group *Xuan Fu Hua* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xuan Fu Hua* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
He Shou Wu	8.46	–
Yin Chen Hao	10.85	–
Sang Ji Shend	13.64	–
Gu Sui Bu	15.81	–
Jing Jie	16.69	–
Dan Zhu Ye	17.12	–
Sang Bai Pi	18.59	–
He Huan Pi	21.75	–
Mang Xiao	23.98	–
Jin Qian Cao	24.86	–
Sha Ren	24.93	–
Du Zhong	25.90	–
Gou Teng	27.06	–
Wu Zhu Yu	27.12	–
Nü Zhen Zi	29.04	–
Ge Gen	32.64	–
Zhi Shi	33.47	–
Bu Gu Zhi	33.68	–
Chen Pi	33.97	–
Sang Ye	34.60	–
Hong Jing Tian	36.65	–
Yi Mu Cao	36.80	–
Bai Hua She She Cao	38.58	–
Hu Zhang	39.00	–
Ding Xiang	39.64	–
Wu Jia Pi	40.79	–
Xia Ku Cao	40.97	–
Xiao Hui Xiang	42.41	–
Tu Si Zi	44.24	–
Jiang Huang	45.97	–
Xuan Shen	46.76	–
Wu Yao	48.51	–
Xin Yi	49.92	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xuan Fu Hua* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62901	62901	0.00	9.08
62902	62902	0.00	8.46

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Xuan Shen**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60095-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Xuan Shen; Scrophulariae ningpoensis radix

Special notes

When selecting the *Xuan Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Xuan Shen	2	0	3

Second-stage model

For differentiation of the substance/substance group *Xuan Shen* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Xuan Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xuan Shen	G223HS092SH1	62553	40	from supplier
PhytoComm	Xuan Shen	G223HS092SH1	62554	40	from supplier
PhytoComm	Xuan Shen	G223HS092ST1	62787	40	from supplier
PhytoComm	Xuan Shen	G223HS092ST1	62788	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Xuan Shen*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Xuan Shen*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Xuan Shen	G223HS092SH1	62553†	20
PhytoComm	Xuan Shen	G223HS092SH1	62554†	20
PhytoComm	Xuan Shen	G223HS092ST1	62787†	20
PhytoComm	Xuan Shen	G223HS092ST1	62788†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Xuan Shen*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Euro OTC	Xuan Shen	H0537021	1
Phytocomm	Xuan Shen	G223H0537421	1
Phytocomm	Xuan Shen	g22h0537322	1

- 854 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Xuan Shen* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Xuan Shen* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	3	854

The substance/substance group *Xuan Shen* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Xuan Shen* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Mang Xiao	20.16	–
Jin Qian Cao	26.08	–
Du Zhong	26.33	–
Jing Jie	27.61	–
Zhi Shi	29.21	–
Wu Jia Pi	30.48	–
Xuan Fu Hua	34.09	–
Yi Mu Cao	34.57	–
Jiu Da Huang	35.19	–
Sang Ye	36.29	–
Ding Xiang	36.78	–
Guang Huo Xiang	37.17	–
Xia Ku Cao	37.44	–
Hong Jing Tian	42.21	–
Sang Ji Shend	42.42	–
Gou Teng	43.84	–
Yin Chen Hao	43.95	–
Yu Xing Cao	44.72	–
Nü Zhen Zi	45.04	–
Xin Yi	46.86	–
Dan Zhu Ye	47.38	–
Hua Shi	48.64	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Xuan Shen* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62553	62553	0.00	22.21
62554	62554	0.00	22.21
62787	62787	0.00	20.16
62788	62788	0.00	20.18

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yan Hu Suo**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60439-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yan Hu Suo; Corydalis rhizoma

Special notes

When selecting the *Yan Hu Suo* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yan Hu Suo	3	0	4

Second-stage model

For differentiation of the substance/substance group *Yan Hu Suo* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yan Hu Suo*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yan Hu Suo	G081H0805821	62371	40	from supplier
PhytoComm	Yan Hu Suo	G081H0805821	62372	40	from supplier
PhytoComm	Yan Hu Suo	G081HS161SG1	62677	40	from supplier
PhytoComm	Yan Hu Suo	G081HS161SG1	62678	40	from supplier
PhytoComm	Yan Hu Suo	G081H0805021	62999	40	from supplier
PhytoComm	Yan Hu Suo	G081H0805021	63000	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Yan Hu Suo*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Yan Hu Suo*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yan Hu Suo	G081H0805821	62371†	20
PhytoComm	Yan Hu Suo	G081H0805821	62372†	20
PhytoComm	Yan Hu Suo	G081HS161SG1	62677†	20
PhytoComm	Yan Hu Suo	G081HS161SG1	62678†	20
PhytoComm	Yan Hu Suo	G081H0805021	62999†	20
PhytoComm	Yan Hu Suo	G081H0805021	63000†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 6 *Apo-Ident* customers from 4 batches from the substance/substance group *Yan Hu Suo*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yan Hu Suo	g081f1000301	1
Phytocomm	Yan Hu Suo	G081H0805223	1
PhytoComm	Yan Hu Suo	G081H0805223	1
Phytocomm	Yan Hu Suo	G081H0805521	1
PhytoComm	Yan Hu Suo	G081H0805521	1
PhytoComm	Yan Hu Suo	G081H0805223	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yan Hu Suo* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yan Hu Suo* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	1	120	0	12 356
Type C	0	1	5	851

The substance/substance group *Yan Hu Suo* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	99.9857 % (> 99.9558 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yan Hu Suo* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Hou Po	5.94	–
Fu Pen Zi	6.29	–
She Gan	8.06	–
Gu Sui Bu	8.58	–
Shan Yao	9.05	–
Pi Pa Ye	9.70	–
Fu Zi	9.84	–
Ce Bai Ye	10.55	–
Dan Dou Chi	11.51	–
Jiang Huang	11.54	–
Gan Cao	11.74	–
Chai Hu	11.75	–
Yu Jin	11.90	–
Ren Dong Teng	12.22	–
Che Qian Zi	12.90	–
Tian Hua Fen	13.37	–
Jin Yin Hua	13.54	–
Zhu Ling	13.59	–
Bai Xian Pi	13.89	–
Mu Zei	13.90	–
Guang Huo Xiang	14.19	–
(Bai) Dou Kou	14.79	–
Yin Yang Huo	14.92	–
Dan Shen	14.96	–
Gan Jiang	15.35	–
Ma Huang	15.38	–
Qiang Huo	15.61	–
Zi Su Zi	15.86	–
Di Gu Pi	15.93	–
Ye Jiao Teng	16.54	–
Qing Pi	16.87	–
Shen Qu	17.00	–
Lu Gen	17.10	–
Mi Huan Jun	17.43	–
Bai Shao Yao	17.44	–
Zi Hua Di Ding	17.91	–
Gua Lou	17.93	–
Tu Fu Ling	18.01	–
Ji Li	18.02	–
Ling Zhi	18.26	–
Lian Qiao	18.36	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
He Huan Pi	18.40	–
Mao Dong Qing	18.50	–
Zhi Gan Cao	18.70	–
Ji Xue Teng	18.75	–
Chuan Xiong	18.76	–
Sha Ren	18.77	–
(Shi) Chang Pu	18.90	–
Ban Lan Gen	19.00	–
Jing Jie	19.12	–
(Fen) Bi Xie	19.31	–
Hong Jing Tian	19.42	–
Lian Zi	19.49	–
Sha Shen (Bei)	19.83	–
Ze Xie	19.86	–
Jiao Gu Lan	20.04	–
Xiao Hui Xiang	20.64	–
Zhe Bei Mu	20.74	–
Suan Zao Ren	20.83	–
Cang Er Zi	20.83	–
Du Zhong	20.87	–
Fu Ling	20.91	–
Zhi Ke	20.99	–
Xiang Fu	21.14	–
Niu Bang Zi	21.64	–
Huang Lian	21.95	–
Tao Ren	22.41	–
Ren Shen	22.47	–
Wu Wei Zi	22.55	–
Sang Zhi	22.80	–
Sheng Jiang	22.82	–
Bai Zi Ren	23.40	–
Bai Zhi	23.71	–
Qing Hao	24.14	–
Chuan Lian Zi	24.28	–
Dang Gui	24.69	–
Huo Ma Ren	25.12	–
(Huai) Niu Xi	25.12	–
Gui Zhi	25.49	–
Yi Mu Cao	25.59	–
Bo He	26.09	–
Jie Geng	26.21	–
Wang Bu Liu Xing	26.25	–
Tai Zi Shen	26.28	–
Huang Bai	26.29	–
E Zhu	26.53	–
Cang Zhu	26.80	–
Gou Teng	26.89	–
Yuan Zhi	27.29	–
Pu Gong Ying	27.53	–
Rou Gui	27.69	–
Bai Hua She She Cao	27.92	–
Chuan Mu Tong	28.36	–
Chen Pi	28.81	–
Gou Qi Zi	28.96	–
Ku Shen	29.59	–
Huang Qin	30.07	–
Mang Xiao	30.36	–
Ma Huang Gen	30.54	–
Yu Zhu	31.58	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Xia (Jiang)	31.69	–
Sang Ye	32.31	–
Wu Yao	32.40	–
Lai Fu Zi	32.43	–
Ze Lan	32.44	–
Huang Qi	33.31	–
Bai Jiang Cao	33.65	–
Yi Yi Ren	34.04	–
San Qi	34.27	–
Bai He	34.54	–
Chuan Niu Xi	34.58	–
Chuang Mu Xiang	34.79	–
Zhu Ru	35.44	–
Shan Yu Rou	35.69	–
Xin Yi	36.06	–
Ci Wu Jia	36.59	–
Hong Hua	36.95	–
Xie Bai	37.21	–
Dang Gui Wei	37.40	–
Du Huo	38.76	–
Fu Xiao Mai	38.99	–
Chi Shao (Yao)	39.02	–
Xu Duan	39.18	–
Xi Xian Cao	39.43	–
Qin Jiao	40.09	–
Sang Ji Shend	40.83	–
Hu Zhang	42.06	–
Bai Zhu	42.17	–
Ju Hua	42.18	–
Fu Shen	43.21	–
Zhi Mu	43.80	–
Mu Dan Pi	45.54	–
Ban Zhi Lian	45.56	–
Fo Shou	46.43	–
Mai Ya	47.16	–
Dong Gua Zi	47.27	–
Yu Xing Cao	48.01	–
Fang Feng	48.18	–
Wu Zhu Yu	48.30	–
Long Yan Rou	48.40	–
Wu Mei	48.68	–
Long Dan (Cao)	49.34	–
He Shou Wu	49.45	–
Nü Zhen Zi	49.95	–
Mu Gua	50.24	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yan Hu Suo* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62371	62371	0.00	11.54
62372	62372	0.00	11.58
62677	62677	0.00	6.29
62678	62678	0.00	5.94
62999	62999	0.00	13.76
63000	63000	0.00	13.65

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ye Jiao Teng**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60068-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ye Jiao Teng; Polygoni multiflori caulis

Special notes

When selecting the *Ye Jiao Teng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ye Jiao Teng	2	0	3

Second-stage model

For differentiation of the substance/substance group *Ye Jiao Teng* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ye Jiao Teng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ye Jiao Teng	G288HS163RM1	62421	40	from supplier
PhytoComm	Ye Jiao Teng	G288HS163RM1	62422	40	from supplier
PhytoComm	Ye Jiao Teng	G288HS163SK1	62625	40	from supplier
PhytoComm	Ye Jiao Teng	G288HS163SK1	62626	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Ye Jiao Teng*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Ye Jiao Teng*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ye Jiao Teng	G288HS163RM1	62421†	20
PhytoComm	Ye Jiao Teng	G288HS163RM1	62422†	20
PhytoComm	Ye Jiao Teng	G288HS163SK1	62625†	20
PhytoComm	Ye Jiao Teng	G288HS163SK1	62626†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Ye Jiao Teng*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ye Jiao Teng	g228h0085223	1
Phytocomm	Ye Jiao Teng	G288H0885223	1
Phytocomm	Ye Jiao Teng	G288H0885522	2
PhytoComm	Ye Jiao Teng	G288H0885522	2

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ye Jiao Teng* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ye Jiao Teng* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	2	158	2	24 438
Type B	1	76	4	12 396
Type C	0	0	6	851

The substance/substance group *Ye Jiao Teng* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	98.7500 % (> 96.8750 %)
Type B	99.9929 % (> 99.9630 %)	95.0000 % (> 91.2500 %)
Type C	100.0000 % (> 98.8258 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ye Jiao Teng* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ji Xue Teng	4.97	–
Bai Xian Pi	6.10	–
Lu Gen	6.16	–
Tu Fu Ling	6.95	–
Dan Dou Chi	7.30	–
Rou Gui	8.96	–
Yin Yang Huo	9.17	–
He Huan Pi	10.06	–
Ren Dong Teng	10.27	–
(Fen) Bi Xie	10.98	–
Fu Zi	11.71	–
Fo Shou	12.23	–
Dan Shen	12.51	–
Ma Huang Gen	12.72	–
Ling Zhi	12.82	–
Chai Hu	12.89	–
Guang Huo Xiang	13.08	–
She Gan	13.18	–
Gou Teng	13.76	–
Bai Shao Yao	14.12	–
Mu Zei	14.34	–
Huo Ma Ren	14.37	–
Mao Dong Qing	14.50	–
Suan Zao Ren	14.55	–
Jing Jie	14.73	–
Gui Zhi	15.03	–
Tai Zi Shen	15.10	–
Zhu Ling	15.22	–
Shen Qu	15.49	–
Lian Zi	15.53	–
Yu Jin	15.74	–
Ma Huang	15.96	–
Tao Ren	16.17	–
Ce Bai Ye	16.41	–
Gu Sui Bu	16.66	–
Hou Po	17.40	–
Yan Hu Suo	17.47	–
Sha Ren	18.10	–
Fu Ling	18.11	–
Qiang Huo	18.85	–
Zi Hua Di Ding	19.13	–
Sheng Jiang	19.24	–
Jin Yin Hua	19.25	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Cao	19.53	–
Fu Pen Zi	20.09	–
Che Qian Zi	20.28	–
Ci Wu Jia	20.46	–
Pi Pa Ye	21.04	–
Chuan Xiong	21.05	–
Shan Yao	21.43	–
Zhe Bei Mu	21.53	–
Yi Yi Ren	21.67	–
Ban Xia (Jiang)	21.88	–
Zhi Ke	22.30	–
Fu Shen	23.08	–
Bo He	23.40	–
Lian Qiao	23.70	–
Ze Xie	23.87	–
Du Zhong	23.92	–
Huang Bai	24.28	–
Lai Fu Zi	24.35	–
Gua Lou	24.38	–
Hong Jing Tian	24.38	–
Ji Li	24.51	–
Di Gu Pi	24.59	–
Huang Lian	25.25	–
Tian Hua Fen	25.34	–
Chen Pi	25.92	–
Yu Zhu	26.11	–
Cang Zhu	27.03	–
Jiao Gu Lan	27.67	–
Ban Lan Gen	28.35	–
Bai Zi Ren	28.98	–
Fu Xiao Mai	29.57	–
Yuan Zhi	30.25	–
Zhi Gan Cao	30.66	–
Mang Xiao	30.77	–
Dang Gui	31.87	–
Huang Qin	31.91	–
Wu Wei Zi	32.13	–
Chuang Mu Xiang	32.26	–
Dang Gui Wei	32.59	–
Qing Pi	32.74	–
Jie Geng	32.96	–
Zhu Ru	33.45	–
Ban Zhi Lian	33.81	–
Ren Shen	34.11	–
Gou Qi Zi	35.22	–
Bai Jiang Cao	35.83	–
Chi Shao (Yao)	36.86	–
Long Yan Rou	36.91	–
Zhi Mu	37.12	–
Sang Zhi	37.46	–
San Qi	39.34	–
Qing Hao	39.60	–
Shan Yu Rou	40.28	–
Hong Hua	40.47	–
Pu Gong Ying	42.61	–
Mu Gua	43.68	–
Bai Zhu	48.08	–
Ku Shen	48.52	–
Ze Lan	48.62	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xie Bai	50.17	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ye Jiao Teng* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62421	62421	0.00	4.97
62422	62422	0.00	5.35
62625	62625	0.00	6.30
62626	62626	0.00	6.10

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yi Mu Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60105-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yi Mu Cao; Leonuri heterophylli herba

Special notes

When selecting the *Yi Mu Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yi Mu Cao	3	0	4

Second-stage model

For differentiation of the substance/substance group *Yi Mu Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yi Mu Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yi Mu Cao	G134H1050821	62409	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050821	62410	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050822	62541	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050822	62542	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050922	62877	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050922	62878	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Yi Mu Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Yi Mu Cao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yi Mu Cao	G134H1050821	62409†	20
PhytoComm	Yi Mu Cao	G134H1050821	62410†	20
PhytoComm	Yi Mu Cao	G134H1050822	62541†	20
PhytoComm	Yi Mu Cao	G134H1050822	62542†	20
PhytoComm	Yi Mu Cao	G134H1050922	62877†	20
PhytoComm	Yi Mu Cao	G134H1050922	62878†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 7 spectra from 5 *Apo-Ident* customers from 4 batches from the substance/substance group *Yi Mu Cao*.
- Among them are spectra of independent samples from 4 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yi Mu Cao	g134h1050024	1
PhytoComm	Yi Mu Cao	G134H1050321	1
Phytocomm	Yi Mu Cao	G134H1050521	1
PhytoComm	Yi Mu Cao	G134H1050521	3
PhytoComm	Yi Mu Cao	G134H1050621	1

- 850 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yi Mu Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yi Mu Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	120	0	12 357
Type C	3	1	6	847

The substance/substance group *Yi Mu Cao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	100.0000 % (> 95.0000 %)
Type C	99.5736 % (> 98.9864 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yi Mu Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Nü Zhen Zi	16.87	–
Cang Er Zi	21.64	–
Xiang Fu	21.90	–
Sha Ren	22.44	–
Jiang Huang	23.54	–
Gu Sui Bu	24.64	–
Xiao Hui Xiang	24.93	–
He Huan Pi	26.81	–
Mang Xiao	27.35	–
Hu Zhang	27.68	–
Ju Hua	27.73	–
Sang Ye	27.86	–
Rou Cong Rong	28.92	–
Gou Teng	29.07	–
Zhi Shi	30.00	–
Ze Lan	30.56	–
Han Lian Cao	30.75	–
Shan Yao	30.83	–
Gua Lou	31.20	–
Bai Hua She She Cao	31.58	–
Shan Yu Rou	31.62	–
Ku Shen	31.74	–
Xu Duan	31.85	–
Xin Yi	31.95	–
Shen Qu	32.55	–
Xi Xian Cao	33.04	–
Chen Pi	33.09	–
Xia Ku Cao	33.20	–
Pu Gong Ying	33.37	–
Wu Zhu Yu	34.12	–
Jing Jie	34.76	–
Sang Ji Shend	35.12	–
Wu Yao	35.42	–
(Bai) Dou Kou	36.58	–
Sang Bai Pi	37.62	–
Dan Zhu Ye	39.18	–
Yu Xing Cao	39.18	–
Niu Bang Zi	39.79	–
Yin Chen Hao	40.32	–
Du Zhong	40.73	–
Gan Cao	41.68	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Qiao	42.46	–
(Shi) Chang Pu	42.86	–
Wu Wei Zi	43.16	–
Huang Lian	43.85	–
Bu Gu Zhi	44.05	–
Tu Fu Ling	45.53	–
Chuan Niu Xi	46.18	–
Hong Jing Tian	46.80	–
Di Gu Pi	46.84	–
Wu Mei	46.89	–
Suan Zao Ren	46.97	–
Yan Hu Suo	47.35	–
Ji Li	47.70	–
Qiang Huo	48.47	–
Fang Feng	49.03	–
He Shou Wu	49.11	–
Qing Pi	49.85	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yi Mu Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62409	62409	0.00	21.64
62410	62410	0.00	21.85
62541	62541	0.00	16.87
62542	62542	0.00	17.21
62877	62877	0.00	21.90
62878	62878	0.00	22.47

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yi Yi Ren**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60174-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yi Yi Ren; Coicis semen

Special notes

When selecting the *Yi Yi Ren* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yi Yi Ren	3	0	2

Second-stage model

For differentiation of the substance/substance group *Yi Yi Ren* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yi Yi Ren*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yi Yi Ren	G078HS336RL1	62279	40	from supplier
PhytoComm	Yi Yi Ren	G078HS336RL1	62280	40	from supplier
PhytoComm	Yi Yi Ren	G078HS336SH1	62675	40	from supplier
PhytoComm	Yi Yi Ren	G078HS336SH1	62676	40	from supplier
PhytoComm	Yi Yi Ren	G078H1701022	63011	40	from supplier
PhytoComm	Yi Yi Ren	G078H1701022	63012	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Yi Yi Ren*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Yi Yi Ren*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yi Yi Ren	G078HS336RL1	62279†	20
PhytoComm	Yi Yi Ren	G078HS336RL1	62280†	20
PhytoComm	Yi Yi Ren	G078HS336SH1	62675†	20
PhytoComm	Yi Yi Ren	G078HS336SH1	62676†	20
PhytoComm	Yi Yi Ren	G078H1701022	63011†	20
PhytoComm	Yi Yi Ren	G078H1701022	63012†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 11 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Yi Yi Ren*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yi Yi Ren	g078h1701323	1
Phytocomm	Yi Yi Ren	G078H1701323	1
PhytoComm	Yi Yi Ren	G078H1701323	6
Phytocomm	Yi Yi Ren	G078H1701522	1
PhytoComm	Yi Yi Ren	G078H1701522	2

- 846 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yi Yi Ren* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yi Yi Ren* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	11	233	7	24 349
Type B	17	116	4	12 340
Type C	0	4	7	846

The substance/substance group *Yi Yi Ren* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9690 % (> 99.9540 %)	97.0833 % (> 95.8333 %)
Type B	99.9119 % (> 99.8820 %)	96.6667 % (> 94.1667 %)
Type C	100.0000 % (> 98.8249 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yi Yi Ren* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Fu Xiao Mai	2.00	–
Zhu Ru	3.69	–
Sheng Jiang	5.48	–
Rou Gui	6.13	–
Fu Ling	7.64	–
Gui Zhi	8.09	–
Di Gu Pi	8.40	–
Fu Shen	9.71	–
Ci Wu Jia	9.76	–
He Huan Pi	10.77	–
Ma Huang Gen	11.93	–
Ji Li	12.61	–
Lian Zi	12.88	–
Ban Xia (Jiang)	12.92	–
Shen Qu	13.05	–
Tao Ren	13.08	–
Bai Shao Yao	14.22	–
Bai Zi Ren	14.65	–
Gou Teng	14.71	–
Ling Zhi	15.12	–
Ji Xue Teng	15.15	–
Mu Zei	16.12	–
Lai Fu Zi	17.04	–
Zhe Bei Mu	17.66	–
Bai Xian Pi	18.64	–
Gua Lou	20.26	–
Huo Ma Ren	20.36	–
Fo Shou	20.89	–
Ye Jiao Teng	20.96	–
Tian Hua Fen	21.43	–
Yuan Zhi	22.01	–
Tai Zi Shen	22.51	–
Shan Yao	22.59	–
Gu Sui Bu	22.83	–
Yu Jin	23.35	–
Pi Pa Ye	25.11	–
Tu Fu Ling	26.98	–
Ren Dong Teng	27.70	–
Lu Gen	27.86	–
Zhu Ling	28.18	–
Lian Qiao	30.06	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ze Xie	30.21	–
Suan Zao Ren	31.39	–
Chuan Xiong	31.58	–
Yan Hu Suo	33.72	–
Mang Xiao	33.81	–
Chen Pi	33.99	–
Shan Yu Rou	34.57	–
(Fen) Bi Xie	35.16	–
Ren Shen	35.22	–
Ce Bai Ye	36.15	–
Fu Zi	37.00	–
Zhi Mu	37.30	–
Dan Dou Chi	37.36	–
San Qi	37.59	–
Cang Zhu	37.78	–
She Gan	37.82	–
Jin Yin Hua	37.89	–
Jiao Gu Lan	37.94	–
Dang Gui Wei	38.19	–
Zhi Ke	38.45	–
Jie Geng	38.59	–
Chai Hu	39.65	–
Fu Pen Zi	39.68	–
Mao Dong Qing	39.75	–
Dan Shen	40.70	–
Wu Wei Zi	42.08	–
Yin Yang Huo	42.73	–
Huang Qin	42.74	–
Yu Zhu	43.31	–
Dang Gui	43.62	–
Che Qian Zi	45.20	–
Hou Po	45.60	–
Mu Gua	48.17	–
Ban Lan Gen	49.40	–
Hong Jing Tian	49.89	–
Guang Huo Xiang	50.58	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yi Yi Ren* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62279	62279	0.00	3.76
62280	62280	0.00	4.58
62675	62675	0.00	2.05
62676	62676	0.00	2.00
63011	63011	0.00	46.24
63012	63012	0.00	46.24

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yin Chen Hao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60163-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yin Chen Hao; Artemisiae scopariae herba

Special notes

When selecting the *Yin Chen Hao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yin Chen Hao	2	0	1

Second-stage model

For differentiation of the substance/substance group *Yin Chen Hao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yin Chen Hao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yin Chen Hao	G034H1030921	62971	40	from supplier
PhytoComm	Yin Chen Hao	G034H1030921	62972	40	from supplier
PhytoComm	Yin Chen Hao	G034HS213TN1	63025	40	from supplier
PhytoComm	Yin Chen Hao	G034HS213TN1	63026	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Yin Chen Hao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Yin Chen Hao*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yin Chen Hao	G034H1030921	62971†	20
PhytoComm	Yin Chen Hao	G034H1030921	62972†	20
PhytoComm	Yin Chen Hao	G034HS213TN1	63025†	20
PhytoComm	Yin Chen Hao	G034HS213TN1	63026†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Yin Chen Hao*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Fagron	Yin Chen Hao	G034H1030421	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yin Chen Hao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yin Chen Hao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	1	80	0	12 396
Type C	0	0	2	855

The substance/substance group *Yin Chen Hao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9952 % (> 99.9654 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yin Chen Hao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Sang Ye	5.99	–
Zhi Shi	8.36	–
Pu Gong Ying	8.59	–
Ze Lan	9.74	–
Dan Zhu Ye	9.84	–
Xin Yi	11.81	–
Xi Xian Cao	12.13	–
Nü Zhen Zi	12.18	–
Xuan Fu Hua	12.69	–
Gu Sui Bu	12.90	–
Sang Bai Pi	13.23	–
Tu Fu Ling	13.46	–
Sang Ji Shend	14.38	–
Shan Yao	14.50	–
Sha Ren	16.27	–
He Shou Wu	16.43	–
Huang Lian	16.60	–
Bu Gu Zhi	17.31	–
Bai Jiang Cao	19.13	–
He Huan Pi	20.81	–
Qiang Huo	21.80	–
Du Zhong	22.21	–
Dan Shen	22.83	–
Fu Zi	23.23	–
Jing Jie	23.62	–
Han Lian Cao	24.25	–
Hu Zhang	24.34	–
Jiang Huang	24.42	–
Wu Mei	24.58	–
Wu Zhu Yu	24.77	–
Mang Xiao	25.84	–
Yu Jin	26.30	–
Zhi Ke	26.57	–
Bo He	26.83	–
Huang Bai	27.18	–
Yu Xing Cao	27.19	–
Tu Si Zi	27.71	–
Zi Hua Di Ding	28.19	–
Mao Dong Qing	28.33	–
Hong Jing Tian	28.66	–
Qing Pi	29.24	–
Chai Hu	29.44	–
Hou Po	29.78	–
Jin Qian Cao	30.68	–
Chen Pi	31.16	–
Bai Hua She She Cao	31.45	–
Jiao Gu Lan	31.47	–
Che Qian Zi	32.00	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Dou Chi	32.46	–
Qing Hao	32.77	–
Wu Wei Zi	32.89	–
Cang Er Zi	33.67	–
Yin Yang Huo	33.67	–
Ge Gen	33.76	–
Yi Mu Cao	34.61	–
Fu Pen Zi	35.00	–
Ye Jiao Teng	35.75	–
Ma Huang	37.09	–
Chuang Mu Xiang	37.47	–
(Bai) Dou Kou	37.50	–
Wang Bu Liu Xing	37.65	–
Jin Yin Hua	37.96	–
Ce Bai Ye	38.01	–
Yan Hu Suo	38.02	–
Hong Hua	38.50	–
Wu Yao	38.61	–
Zi Su Zi	38.71	–
Ji Li	38.96	–
Di Gu Pi	38.99	–
Bai Xian Pi	39.87	–
Xiang Fu	39.93	–
Wu Jia Pi	41.79	–
Gua Lou	41.94	–
Niu Bang Zi	41.96	–
Xia Ku Cao	42.42	–
Chuan Xiong	43.00	–
Suan Zao Ren	43.36	–
Gou Teng	43.83	–
Gan Cao	44.04	–
Ling Zhi	44.63	–
Pi Pa Ye	44.64	–
(Shi) Chang Pu	45.02	–
E Zhu	45.26	–
Mu Dan Pi	45.82	–
Lian Qiao	46.30	–
Shan Yu Rou	47.04	–
Lian Zi	47.67	–
Ren Dong Teng	47.97	–
Bai Shao Yao	48.05	–
Chi Shao (Yao)	49.28	–
Zhe Bei Mu	49.48	–
Hua Shi	49.70	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yin Chen Hao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62971	62971	0.00	9.84
62972	62972	0.00	10.56
63025	63025	0.00	6.18
63026	63026	0.00	5.99

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yin Yang Huo**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60228-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yin Yang Huo; Epimedii herba

Special notes

When selecting the *Yin Yang Huo* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yin Yang Huo	1	0	1

Second-stage model

For differentiation of the substance/substance group *Yin Yang Huo* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yin Yang Huo*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yin Yang Huo	G098HS241SH1	62683	40	from supplier
PhytoComm	Yin Yang Huo	G098HS241SH1	62684	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Yin Yang Huo*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Yin Yang Huo*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yin Yang Huo	G098HS241SH1	62683†	20
PhytoComm	Yin Yang Huo	G098HS241SH1	62684†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 1 batches from the substance/substance group *Yin Yang Huo*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yin Yang Huo	G098H1114422	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yin Yang Huo* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yin Yang Huo* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	3	80	0	24 517
Type B	1	39	1	12 436
Type C	0	0	2	855

The substance/substance group *Yin Yang Huo* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9905 % (> 99.9755 %)	100.0000 % (> 92.5000 %)
Type B	99.9952 % (> 99.9655 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yin Yang Huo* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Xian Pi	4.16	–
Jing Jie	4.43	–
Mao Dong Qing	5.21	–
Dan Dou Chi	6.44	–
Hou Po	6.84	–
Ling Zhi	6.84	–
Yu Jin	8.13	–
Ye Jiao Teng	8.29	–
Chai Hu	8.64	–
Che Qian Zi	8.85	–
Qiang Huo	9.24	–
Fu Zi	9.28	–
Ce Bai Ye	9.91	–
Zi Hua Di Ding	10.61	–
Sha Ren	11.73	–
Ren Dong Teng	11.88	–
Shen Qu	12.01	–
Yan Hu Suo	12.15	–
Tu Fu Ling	12.44	–
Zhu Ling	13.09	–
Du Zhong	13.97	–
Lian Zi	14.12	–
Shan Yao	14.47	–
Dan Shen	14.55	–
Fu Pen Zi	14.62	–
Guang Huo Xiang	14.71	–
Bo He	14.84	–
Ji Xue Teng	16.09	–
Pi Pa Ye	16.40	–
Huang Bai	16.73	–
Ma Huang	17.24	–
Chuan Xiong	17.44	–
Gu Sui Bu	18.00	–
(Fen) Bi Xie	18.75	–
Ji Li	18.93	–
Rou Gui	19.83	–
Qing Pi	19.97	–
Jin Yin Hua	20.01	–
Qing Hao	20.08	–
Tian Hua Fen	20.31	–
Ma Huang Gen	20.53	–
Gan Cao	20.63	–
He Huan Pi	20.88	–
Hong Jing Tian	20.93	–
Suan Zao Ren	21.83	–
Lian Qiao	21.91	–
Gou Teng	21.94	–
Huang Lian	22.88	–
Zhi Ke	22.99	–
Lu Gen	23.07	–
Jiao Gu Lan	23.10	–
Bai Shao Yao	23.12	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
She Gan	23.69	–
Sheng Jiang	24.20	–
Bai Jiang Cao	24.30	–
Gua Lou	25.10	–
Ban Lan Gen	25.49	–
Huo Ma Ren	25.51	–
Fu Ling	25.73	–
Di Gu Pi	25.78	–
Mu Zei	25.90	–
Gui Zhi	26.58	–
Yi Yi Ren	26.90	–
Wu Wei Zi	27.08	–
Zhi Gan Cao	29.46	–
Cang Zhu	29.54	–
Yuan Zhi	29.95	–
Chuang Mu Xiang	29.99	–
Ren Shen	30.69	–
Zhe Bei Mu	31.15	–
Tao Ren	31.49	–
Mang Xiao	32.00	–
Gou Qi Zi	32.95	–
Tai Zi Shen	33.50	–
Pu Gong Ying	34.82	–
Ban Zhi Lian	36.57	–
Chen Pi	36.68	–
Bai Zi Ren	36.69	–
Ban Xia (Jiang)	37.39	–
Ci Wu Jia	37.84	–
Fu Shen	38.25	–
Fo Shou	38.48	–
Fu Xiao Mai	38.55	–
Chi Shao (Yao)	38.81	–
Sang Zhi	38.96	–
Hong Hua	39.13	–
Ze Lan	39.39	–
Yu Zhu	39.67	–
Huang Qin	40.64	–
Ze Xie	40.74	–
Zhu Ru	41.15	–
Jie Geng	41.71	–
Dang Gui Wei	43.67	–
Dang Gui	44.23	–
Cang Er Zi	44.63	–
Sang Ye	44.95	–
Xie Bai	44.98	–
Xi Xian Cao	45.37	–
E Zhu	46.90	–
San Qi	46.93	–
Shan Yu Rou	47.42	–
Jiang Huang	47.72	–
Long Yan Rou	48.04	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yin Yang Huo* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62683	62683	0.00	4.17
62684	62684	0.00	4.16

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yu Jin**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60171-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yu Jin; Curcumae tuber

Special notes

When selecting the *Yu Jin* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yu Jin	1	0	0

Second-stage model

For differentiation of the substance/substance group *Yu Jin* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yu Jin*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yu Jin	G084HS357RN1	62277	40	from supplier
PhytoComm	Yu Jin	G084HS357RN1	62278	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Yu Jin*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Yu Jin*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yu Jin	G084HS357RN1	62277†	20
PhytoComm	Yu Jin	G084HS357RN1	62278†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Yu Jin*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yu Jin* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yu Jin* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	3	40	0	12 434
Type C	0	0	0	857

The substance/substance group *Yu Jin* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9833 % (> 99.9536 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8345 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yu Jin* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Ling Zhi	5.64	–
Mao Dong Qing	6.30	–
Hou Po	6.84	–
Yin Yang Huo	7.16	–
Qiang Huo	7.35	–
Du Zhong	8.67	–
Jing Jie	8.91	–
Bai Xian Pi	8.99	–
Dan Dou Chi	9.31	–
Ce Bai Ye	9.66	–
Che Qian Zi	10.09	–
Qing Hao	11.76	–
Ye Jiao Teng	12.35	–
Fu Zi	13.34	–
Ren Dong Teng	13.37	–
Yan Hu Suo	13.83	–
Sha Ren	13.88	–
Zi Hua Di Ding	14.07	–
Shen Qu	14.46	–
Shan Yao	14.58	–
Chai Hu	14.73	–
Fu Pen Zi	14.76	–
Dan Shen	15.52	–
Qing Pi	15.66	–
Zhu Ling	16.44	–
Guang Huo Xiang	16.99	–
Bo He	17.01	–
Jiao Gu Lan	17.99	–
Pi Pa Ye	18.04	–
Ji Li	18.89	–
Lian Zi	18.99	–
Huang Bai	19.41	–
Chuan Xiong	19.47	–
Hong Jing Tian	19.60	–
Ji Xue Teng	20.05	–
Lian Qiao	20.08	–
He Huan Pi	20.93	–
Gu Sui Bu	21.22	–
Tian Hua Fen	22.25	–
Zhi Ke	22.38	–
Huang Lian	22.52	–
Ma Huang	22.60	–
Jin Yin Hua	22.65	–
Bai Jiang Cao	23.43	–
Suan Zao Ren	23.92	–
Wu Wei Zi	24.11	–
She Gan	24.35	–
Gou Teng	25.37	–
(Fen) Bi Xie	25.96	–
Bai Shao Yao	26.00	–
Tu Fu Ling	26.34	–
Di Gu Pi	26.76	–
Zhe Bei Mu	27.01	–
Gan Cao	27.01	–
Sheng Jiang	27.63	–
Mu Zei	27.80	–
Lu Gen	27.88	–
Ban Lan Gen	28.43	–
Huo Ma Ren	29.69	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gua Lou	30.41	–
Ma Huang Gen	30.45	–
Ren Shen	30.93	–
Yuan Zhi	31.56	–
Chuang Mu Xiang	31.69	–
Fu Ling	31.95	–
Tao Ren	32.00	–
Rou Gui	32.33	–
Cang Zhu	32.68	–
Mang Xiao	32.91	–
Pu Gong Ying	33.74	–
Yi Yi Ren	33.82	–
Tai Zi Shen	34.30	–
Zhi Gan Cao	34.90	–
Gui Zhi	35.57	–
Xi Xian Cao	35.76	–
Ze Lan	36.52	–
Gou Qi Zi	37.05	–
Sang Ye	38.21	–
Bai Zi Ren	40.40	–
Chen Pi	40.45	–
Ban Zhi Lian	40.82	–
Cang Er Zi	41.18	–
Jie Geng	41.59	–
Ze Xie	42.09	–
Nü Zhen Zi	42.35	–
Ci Wu Jia	42.43	–
Hong Hua	42.91	–
Ban Xia (Jiang)	43.00	–
Sang Zhi	43.25	–
Xie Bai	43.49	–
Yu Zhu	43.73	–
Chi Shao (Yao)	43.79	–
Huang Qin	43.96	–
Jiang Huang	44.01	–
Zhu Ru	44.18	–
Dang Gui	44.51	–
E Zhu	44.66	–
Shan Yu Rou	45.86	–
San Qi	46.09	–
Fu Xiao Mai	46.84	–
Wu Mei	47.01	–
Fu Shen	47.18	–
Ku Shen	48.15	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yu Jin* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62277	62277	0.00	5.64
62278	62278	0.00	5.97

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yu Xing Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60606-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yu Xing Cao; Houttuyniae herba

Special notes

When selecting the *Yu Xing Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yu Xing Cao	2	0	3

Second-stage model

For differentiation of the substance/substance group *Yu Xing Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yu Xing Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yu Xing Cao	G124H1149823	62583	40	from supplier
PhytoComm	Yu Xing Cao	G124H1149823	62584	40	from supplier
PhytoComm	Yu Xing Cao	G124HS244TL1	62889	40	from supplier
PhytoComm	Yu Xing Cao	G124HS244TL1	62890	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Yu Xing Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Yu Xing Cao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yu Xing Cao	G124H1149823	62583†	20
PhytoComm	Yu Xing Cao	G124H1149823	62584†	20
PhytoComm	Yu Xing Cao	G124HS244TL1	62889†	20
PhytoComm	Yu Xing Cao	G124HS244TL1	62890†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 3 *Apo-Ident* customers from 4 batches from the substance/substance group *Yu Xing Cao*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yu Xing Cao	g124h1149122	1
Phytocomm	Yu Xing Cao	g124h1149222	1
Phytocomm	Yu Xing Cao	G124H1149222	1
Phytocomm	Yu Xing Cao	G124H1149423	2

- 852 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yu Xing Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yu Xing Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	0	0	5	852

The substance/substance group *Yu Xing Cao* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yu Xing Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Han Lian Cao	8.87	–
Xi Xian Cao	9.75	–
Hu Zhang	11.92	–
Bai Hua She She Cao	12.59	–
Xin Yi	12.81	–
(Sheng) Di Huang	13.19	–
Pu Gong Ying	14.67	–
Wu Jia Pi	15.17	–
Sang Ye	15.25	–
Ze Lan	16.02	–
Tu Fu Ling	17.13	–
Huang Lian	17.57	–
Xian Mao	17.77	–
Jiang Huang	17.94	–
Zhi Shi	19.48	–
Xia Ku Cao	21.14	–
Shan Yao	21.40	–
Mang Xiao	21.94	–
Nü Zhen Zi	23.38	–
Shu Di (Huang)	24.07	–
Xiang Fu	25.31	–
Yin Chen Hao	26.22	–
Sha Ren	27.30	–
Jing Jie	29.05	–
Yi Mu Cao	31.89	–
He Huan Pi	32.98	–
Du Zhong	35.63	–
Qiang Huo	36.27	–
Fu Zi	37.91	–
Yu Jin	38.77	–
Wu Mei	39.84	–
Yan Hu Suo	39.85	–
Sang Ji Shend	40.29	–
Bai Jiang Cao	40.48	–
Dan Shen	40.82	–
Bo He	40.92	–
Dan Zhu Ye	40.93	–
Rou Cong Rong	41.22	–
Ding Xiang	41.36	–
Gu Sui Bu	42.27	–
Cang Er Zi	42.74	–
Che Qian Zi	43.13	–
Mao Dong Qing	44.27	–
Huang Bai	44.32	–
Guang Huo Xiang	44.79	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	44.96	–
E Zhu	45.41	–
Hua Shi	45.71	–
Hou Po	46.19	–
Gua Lou	46.36	–
Jiao Gu Lan	46.48	–
Hong Jing Tian	47.26	–
Wang Bu Liu Xing	47.62	–
Dan Dou Chi	48.06	–
Ji Li	48.24	–
Gou Teng	48.52	–
Qing Pi	48.53	–
Yin Yang Huo	49.34	–
(Shi) Chang Pu	50.13	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yu Xing Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62583	62583	0.00	13.92
62584	62584	0.00	13.19
62889	62889	0.00	9.38
62890	62890	0.00	8.87

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yu Zhu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60056-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yu Zhu; Polygonati odorati rhizoma

Special notes

When selecting the *Yu Zhu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yu Zhu	1	0	2

Second-stage model

For differentiation of the substance/substance group *Yu Zhu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yu Zhu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yu Zhu	G195HS087SK1	62605	40	from supplier
PhytoComm	Yu Zhu	G195HS087SK1	62606	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Yu Zhu*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Yu Zhu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yu Zhu	G195HS087SK1	62605†	20
PhytoComm	Yu Zhu	G195HS087SK1	62606†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Yu Zhu*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Yu Zhu	G195H0517121	1
Phytocomm	Yu Zhu	G195H0517421	2

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yu Zhu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yu Zhu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	79	1	24 520
Type B	1	36	4	12 436
Type C	0	0	3	854

The substance/substance group *Yu Zhu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	98.7500 % (> 95.0000 %)
Type B	99.9929 % (> 99.9631 %)	90.0000 % (> 82.5000 %)
Type C	100.0000 % (> 98.8280 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yu Zhu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Cang Zhu	2.61	–
Dang Gui Wei	7.93	–
Dang Gui	7.97	–
Gua Lou	8.32	–
Chen Pi	8.52	–
Chuan Xiong	10.83	–
Jin Yin Hua	12.59	–
Gou Qi Zi	12.91	–
Jie Geng	13.06	–
Yuan Zhi	13.81	–
Zhe Bei Mu	13.90	–
Tai Zi Shen	13.90	–
Bai Shao Yao	14.17	–
Pi Pa Ye	14.80	–
Ze Xie	15.07	–
Tian Hua Fen	15.25	–
Long Yan Rou	15.35	–
Mu Zei	15.66	–
San Qi	16.26	–
Fo Shou	16.50	–
Ye Jiao Teng	17.00	–
Chuang Mu Xiang	17.22	–
Ren Dong Teng	17.49	–
Suan Zao Ren	17.53	–
Mu Gua	18.26	–
Lai Fu Zi	18.31	–
Di Gu Pi	18.31	–
Shen Qu	18.40	–
Zhi Mu	18.59	–
Dan Dou Chi	18.63	–
Lian Zi	18.72	–
He Huan Pi	19.51	–
Lian Qiao	19.68	–
Ban Lan Gen	19.74	–
Tao Ren	19.86	–
Hong Jing Tian	19.91	–
Zhi Ke	20.11	–
Gu Sui Bu	20.53	–
Xie Bai	20.57	–
Ren Shen	20.68	–
Ling Zhi	20.73	–
Ma Huang	20.94	–
Bai Zi Ren	21.02	–
Shan Yao	22.15	–
Fu Pen Zi	22.23	–
Lu Gen	22.42	–
Ci Wu Jia	22.48	–
Ji Li	22.71	–
Rou Gui	22.82	–
Ku Shen	22.89	–
Jiao Gu Lan	23.02	–
Gou Teng	23.22	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Guang Huo Xiang	23.41	–
Dan Shen	23.83	–
Bai Xian Pi	24.24	–
Sheng Jiang	24.40	–
Gui Zhi	24.48	–
Tu Fu Ling	24.82	–
Shan Yu Rou	25.17	–
Hou Po	25.53	–
Ji Xue Teng	25.73	–
Fu Zi	26.20	–
Huo Ma Ren	27.89	–
Zhu Ru	29.38	–
Huang Qin	29.39	–
Fu Ling	29.76	–
Gan Cao	29.85	–
She Gan	29.96	–
Yin Yang Huo	29.99	–
Wu Wei Zi	30.40	–
Ban Xia (Jiang)	30.47	–
Chai Hu	30.53	–
Mang Xiao	31.07	–
Mao Dong Qing	31.18	–
Ma Huang Gen	32.10	–
Fu Xiao Mai	32.15	–
Che Qian Zi	32.94	–
Yan Hu Suo	33.55	–
(Fen) Bi Xie	34.10	–
Ce Bai Ye	34.41	–
Zi Hua Di Ding	35.51	–
Yu Jin	35.94	–
Mai Men Dong	36.18	–
Yi Yi Ren	36.52	–
Sang Zhi	36.56	–
Qiang Huo	37.16	–
Zhi Gan Cao	38.98	–
Fu Shen	39.63	–
Chi Shao (Yao)	39.65	–
Bo He	40.86	–
Hong Hua	41.04	–
Zhu Ling	41.05	–
Jing Jie	41.81	–
Huang Bai	44.19	–
Qing Pi	44.62	–
Sha Ren	45.44	–
Huang Lian	46.08	–
Cang Er Zi	47.93	–
Chuan Lian Zi	48.88	–
Huang Qi	49.12	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yu Zhu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62605	62605	0.00	3.29
62606	62606	0.00	2.61

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Yuan Zhi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60100-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Yuan Zhi; Polygalae radix

Special notes

When selecting the *Yuan Zhi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Yuan Zhi	3	0	3

Second-stage model

For differentiation of the substance/substance group *Yuan Zhi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Yuan Zhi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Yuan Zhi	G194HS306SH1	62557	40	from supplier
PhytoComm	Yuan Zhi	G194HS306SH1	62558	40	from supplier
PhytoComm	Yuan Zhi	G194H1406921	62743	40	from supplier
PhytoComm	Yuan Zhi	G194H1406921	62744	40	from supplier
PhytoComm	Yuan Zhi	G194HS306TG1	62845	40	from supplier
PhytoComm	Yuan Zhi	G194HS306TG1	62846	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Yuan Zhi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Yuan Zhi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Yuan Zhi	G194HS306SH1	62557†	20
PhytoComm	Yuan Zhi	G194HS306SH1	62558†	20
PhytoComm	Yuan Zhi	G194H1406921	62743†	20
PhytoComm	Yuan Zhi	G194H1406921	62744†	20
PhytoComm	Yuan Zhi	G194HS306TG1	62845†	20
PhytoComm	Yuan Zhi	G194HS306TG1	62846†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 6 *Apo-Ident* customers from 3 batches from the substance/substance group *Yuan Zhi*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Yuan Zhi	g194h140622	1
Phytocomm	Yuan Zhi	G194H1406222	1
PhytoComm	Yuan Zhi	G194H1406222	1
Phytocomm	Yuan Zhi	G194H1406522	2
PhytoComm	Yuan Zhi	G194H1406522	1

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Yuan Zhi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Yuan Zhi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	3	240	0	24 357
Type B	5	119	1	12 352
Type C	1	2	4	850

The substance/substance group *Yuan Zhi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9786 % (> 99.9635 %)	100.0000 % (> 97.5000 %)
Type B	99.9381 % (> 99.9082 %)	99.1667 % (> 96.6667 %)
Type C	99.9225 % (> 99.3354 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Yuan Zhi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Bai Shao Yao	5.02	–
Zhi Mu	6.28	–
(Huai) Niu Xi	9.44	–
Ren Shen	11.50	–
Jie Geng	11.64	–
Pi Pa Ye	11.67	–
Shan Yao	12.03	–
Ye Jiao Teng	12.38	–
Gua Lou	12.61	–
Shan Yu Rou	12.88	–
Huang Qi	12.93	–
Fo Shou	13.96	–
Zhu Ru	14.44	–
Di Gu Pi	14.53	–
Bai Zhu	15.57	–
San Qi	15.72	–
Ji Li	15.72	–
Dang Gui Wei	16.23	–
Shen Qu	16.91	–
Ku Shen	17.10	–
Fu Xiao Mai	17.32	–
He Huan Pi	17.37	–
Lian Qiao	17.74	–
Lian Zi	18.23	–
Mu Gua	18.24	–
Cang Zhu	18.27	–
Tian Hua Fen	18.59	–
Gou Qi Zi	18.73	–
Dang Gui	19.15	–
Ci Wu Jia	19.58	–
Chi Shao (Yao)	19.68	–
Rou Gui	19.68	–
Sang Zhi	19.83	–
Bai Zi Ren	20.13	–
Zhe Bei Mu	20.29	–
Sheng Jiang	20.75	–
Gui Zhi	21.81	–
Zhi Gan Cao	21.93	–
Tai Zi Shen	22.05	–
Mai Men Dong	22.41	–
Lai Fu Zi	22.43	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chuan Lian Zi	22.48	–
Fu Ling	22.56	–
Zi Su Zi	22.86	–
Ren Dong Teng	22.90	–
Ling Zhi	23.23	–
Niu Bang Zi	23.55	–
Chuan Xiong	24.03	–
Mu Zei	24.95	–
Tu Fu Ling	25.14	–
Huo Ma Ren	25.14	–
Xiang Fu	25.67	–
Yi Yi Ren	25.97	–
Jin Yin Hua	26.06	–
Yu Zhu	26.06	–
Jiao Gu Lan	26.08	–
Ban Xia (Jiang)	26.90	–
Tao Ren	26.98	–
Zhi Ke	27.10	–
Gan Cao	27.68	–
Chen Pi	28.21	–
Xie Bai	28.97	–
Mang Xiao	29.13	–
Ban Lan Gen	29.21	–
Gou Teng	29.29	–
Suan Zao Ren	29.35	–
Fu Zi	29.36	–
Ji Xue Teng	29.44	–
Wu Wei Zi	30.41	–
Bai Xian Pi	30.96	–
(Shi) Chang Pu	31.39	–
Ze Xie	31.84	–
Lu Gen	32.00	–
Chuang Mu Xiang	32.30	–
Huang Qin	32.73	–
Hou Po	32.95	–
Fu Pen Zi	33.12	–
Mao Dong Qing	33.51	–
Long Yan Rou	33.68	–
Gu Sui Bu	33.93	–
Long Dan (Cao)	34.62	–
Qin Jiao	34.75	–
Chuan Mu Tong	34.91	–
Cang Er Zi	35.27	–
Bing Lang	35.32	–
Dan Dou Chi	35.39	–
Che Qian Zi	35.71	–
Yu Jin	36.75	–
Dan Shen	36.85	–
Gan Jiang	37.15	–
Du Huo	37.26	–
Bai He	37.71	–
Yan Hu Suo	37.97	–
Bai Zhi	38.26	–
Chuan Niu Xi	38.27	–
Pu Gong Ying	38.75	–
E Zhu	39.71	–
Sha Shen (Bei)	40.03	–
Ma Huang Gen	40.29	–
Dong Gua Zi	41.42	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ce Bai Ye	41.96	–
Qiang Huo	42.26	–
Chai Hu	42.33	–
Yin Yang Huo	42.42	–
Yi Mu Cao	42.76	–
Ju Hua	43.03	–
Fu Shen	43.05	–
Xiao Hui Xiang	43.43	–
Guang Huo Xiang	43.93	–
(Fen) Bi Xie	44.46	–
Zhu Ling	44.48	–
Hong Jing Tian	44.65	–
Wu Yao	45.54	–
Ma Huang	45.74	–
Xu Duan	47.43	–
Ba Ji Tian	47.76	–
She Gan	47.93	–
Zi Hua Di Ding	48.33	–
Bo He	48.39	–
(Bai) Dou Kou	48.65	–
Fang Feng	48.73	–
Sha Ren	49.33	–
Jiang Huang	49.71	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Yuan Zhi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62557	62557	0.00	5.02
62558	62558	0.00	5.32
62743	62743	0.00	9.44
62744	62744	0.00	9.63
62845	62845	0.00	8.39
62846	62846	0.00	8.02

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ze Lan**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60322-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ze Lan; Lycopi lucidi herba

Special notes

When selecting the *Ze Lan* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ze Lan	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ze Lan* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ze Lan*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ze Lan	G268HS321TK1	62813	40	from supplier
PhytoComm	Ze Lan	G268HS321TK1	62814	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ze Lan*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ze Lan*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ze Lan	G268HS321TK1	62813†	20
PhytoComm	Ze Lan	G268HS321TK1	62814†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 4 spectra from 3 *Apo-Ident* customers from 3 batches from the substance/substance group *Ze Lan*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ze Lan	g268h1606221	1
Phytocomm	Ze Lan	G268H1606221	1
PhytoComm	Ze Lan	G268H1606221	1
Phytocomm	Ze Lan	G268H1606421	1

- 853 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ze Lan* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ze Lan* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	2	40	0	12 435
Type C	0	0	4	853

The substance/substance group *Ze Lan* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	99.9881 % (> 99.9584 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8269 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ze Lan* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Xi Xian Cao	7.91	–
Tu Fu Ling	8.87	–
Pu Gong Ying	10.91	–
Yin Chen Hao	11.69	–
Huang Lian	12.38	–
Xin Yi	15.15	–
Sang Ye	18.24	–
Sha Ren	18.28	–
Bai Jiang Cao	18.34	–
Zhi Ke	18.37	–
Shan Yao	19.40	–
Yu Jin	19.49	–
Nü Zhen Zi	20.08	–
Jiao Gu Lan	20.11	–
Du Zhong	20.74	–
Zhi Shi	21.06	–
Bai Hua She She Cao	22.14	–
Dan Shen	22.35	–
Fu Zi	22.62	–
Yu Xing Cao	22.67	–
Jiang Huang	24.60	–
Hu Zhang	25.13	–
Mao Dong Qing	25.67	–
Bo He	25.85	–
Qiang Huo	25.93	–
Che Qian Zi	26.02	–
Huang Bai	27.81	–
Chai Hu	28.91	–
Mang Xiao	29.31	–
Han Lian Cao	29.98	–
Bai Xian Pi	30.01	–
Hou Po	30.53	–
Ye Jiao Teng	30.73	–
Yin Yang Huo	32.65	–
Yan Hu Suo	32.87	–
Zi Hua Di Ding	33.04	–
Wu Mei	33.09	–
Dan Dou Chi	33.27	–
Ling Zhi	34.40	–
Cang Er Zi	34.89	–
Hong Jing Tian	35.30	–
Qing Hao	35.91	–
Wu Wei Zi	35.98	–
Fu Pen Zi	36.04	–
Di Gu Pi	36.05	–
Qing Pi	36.46	–
Pi Pa Ye	37.37	–
Wang Bu Liu Xing	37.78	–
Jin Yin Hua	38.13	–
Chuang Mu Xiang	38.60	–
Ce Bai Ye	38.85	–
He Huan Pi	39.27	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	39.35	–
Yi Mu Cao	39.97	–
Jing Jie	40.17	–
Lian Zi	40.67	–
Ma Huang	40.70	–
Xiang Fu	42.38	–
Chuan Xiong	42.55	–
Zi Su Zi	42.65	–
(Shi) Chang Pu	43.00	–
Suan Zao Ren	43.24	–
Ren Dong Teng	43.66	–
Bai Shao Yao	43.79	–
Ji Xue Teng	44.32	–
Niu Bang Zi	45.64	–
Ji Li	45.94	–
Sha Shen (Bei)	47.02	–
Tian Hua Fen	47.08	–
Gua Lou	47.25	–
Lian Qiao	47.95	–
(Bai) Dou Kou	48.17	–
Gu Sui Bu	48.74	–
Sang Ji Shend	48.89	–
Shen Qu	49.17	–
Ren Shen	49.21	–
Xia Ku Cao	49.24	–
Chi Shao (Yao)	49.48	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ze Lan* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62813	62813	0.00	8.87
62814	62814	0.00	7.91

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Ze Xie**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10004523-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Ze Xie; Alismatis rhizoma

Special notes

When selecting the *Ze Xie* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Ze Xie	1	0	3

Second-stage model

For differentiation of the substance/substance group *Ze Xie* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Ze Xie*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ze Xie	G011HS320TG1	62955	40	from supplier
PhytoComm	Ze Xie	G011HS320TG1	62956	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Ze Xie*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Ze Xie*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Ze Xie	G011HS320TG1	62955†	20
PhytoComm	Ze Xie	G011HS320TG1	62956†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 9 spectra from 5 *Apo-Ident* customers from 4 batches from the substance/substance group *Ze Xie*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Ze Xie	g011h1605122	1
PhytoComm	Ze Xie	G011H1605122	1
Phytocomm	Ze Xie	G011H1605422	1
Phytocomm	Ze Xie	G011H1605521	5
PhytoComm	Ze Xie	G011H1605521	1

- 848 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Ze Xie* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Ze Xie* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	79	1	24 520
Type B	0	37	3	12 437
Type C	0	0	9	848

The substance/substance group *Ze Xie* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9406 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8251 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Ze Xie* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Pi Pa Ye	6.08	–
Gua Lou	6.62	–
Tian Hua Fen	7.83	–
Ren Dong Teng	8.13	–
Bai Shao Yao	8.61	–
Fu Zi	8.75	–
Mu Zei	8.80	–
Shen Qu	8.87	–
Shan Yao	9.02	–
Lian Qiao	9.27	–
Tai Zi Shen	9.51	–
Zhe Bei Mu	9.75	–
Chuan Xiong	10.33	–
Ban Lan Gen	10.40	–
Jin Yin Hua	10.92	–
Ling Zhi	11.91	–
Chen Pi	12.10	–
He Huan Pi	12.19	–
Huo Ma Ren	12.30	–
Lai Fu Zi	12.46	–
Gu Sui Bu	12.51	–
Ye Jiao Teng	12.56	–
Ji Li	12.61	–
Fu Pen Zi	12.89	–
Yuan Zhi	13.40	–
Lian Zi	13.43	–
Lu Gen	13.86	–
Dan Dou Chi	13.97	–
Bai Zi Ren	14.04	–
Yu Zhu	14.40	–
Suan Zao Ren	14.54	–
Tao Ren	14.54	–
Cang Zhu	14.55	–
Hou Po	15.01	–
Gui Zhi	15.47	–
Ji Xue Teng	15.47	–
Jiao Gu Lan	15.65	–
Dang Gui Wei	15.96	–
Di Gu Pi	16.06	–
Sheng Jiang	16.13	–
Ci Wu Jia	16.18	–
Dang Gui	16.40	–
Jie Geng	16.57	–
Dan Shen	16.85	–
Guang Huo Xiang	16.97	–
Gou Teng	17.03	–
Fu Ling	17.59	–
San Qi	18.01	–
Yan Hu Suo	18.78	–
She Gan	19.09	–
Hong Jing Tian	19.11	–
Ce Bai Ye	19.44	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	19.46	–
Ma Huang	19.71	–
Che Qian Zi	19.76	–
Zhu Ru	20.09	–
Chai Hu	20.24	–
Yin Yang Huo	20.28	–
Bai Xian Pi	20.31	–
Gou Qi Zi	20.34	–
Tu Fu Ling	20.55	–
Rou Gui	20.67	–
Ren Shen	20.88	–
Yu Jin	21.09	–
Ku Shen	21.44	–
Fo Shou	21.72	–
(Fen) Bi Xie	21.86	–
Mao Dong Qing	22.00	–
Ban Xia (Jiang)	22.07	–
Wu Wei Zi	22.46	–
Zhu Ling	23.00	–
Zhi Mu	23.02	–
Gan Cao	23.05	–
Qiang Huo	24.73	–
Zhi Gan Cao	25.94	–
Shan Yu Rou	26.00	–
Yi Yi Ren	26.20	–
Fu Xiao Mai	26.47	–
Huang Qin	26.60	–
Ma Huang Gen	27.24	–
Chuang Mu Xiang	27.27	–
Mu Gua	27.57	–
Xie Bai	27.60	–
Long Yan Rou	28.71	–
Zi Hua Di Ding	30.74	–
Qing Pi	30.78	–
Mang Xiao	32.15	–
Jing Jie	32.57	–
Huang Lian	33.96	–
Bo He	34.18	–
Sang Zhi	34.81	–
Chi Shao (Yao)	36.29	–
Sha Ren	36.65	–
Fu Shen	37.60	–
Huang Bai	38.04	–
Du Zhong	38.31	–
Hong Hua	42.35	–
Cang Er Zi	43.42	–
Bai Jiang Cao	49.84	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ze Xie* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested

reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62955	62955	0.00	6.08
62956	62956	0.00	6.55

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhe Bei Mu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50388-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhe Bei Mu; Fritillariae thunbergii bulbus

Special notes

When selecting the *Zhe Bei Mu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhe Bei Mu	2	0	2

Second-stage model

For differentiation of the substance/substance group *Zhe Bei Mu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhe Bei Mu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhe Bei Mu	G108HS138SG1	62687	40	from supplier
PhytoComm	Zhe Bei Mu	G108HS138SG1	62688	40	from supplier
PhytoComm	Zhe Bei Mu	G108HS138TT1	63027	40	from supplier
PhytoComm	Zhe Bei Mu	G108HS138TT1	63028	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Zhe Bei Mu*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Zhe Bei Mu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhe Bei Mu	G108HS138SG1	62687†	20
PhytoComm	Zhe Bei Mu	G108HS138SG1	62688†	20
PhytoComm	Zhe Bei Mu	G108HS138TT1	63027†	20
PhytoComm	Zhe Bei Mu	G108HS138TT1	63028†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 2 *Apo-Ident* customers from 2 batches from the substance/substance group *Zhe Bei Mu*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Zhe Bei Mu	G108H0712321	1
Phytocomm	Zhe Bei Mu	G108H0712422	1

- 855 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhe Bei Mu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhe Bei Mu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	1	77	3	12 396
Type C	0	0	2	855

The substance/substance group *Zhe Bei Mu* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	99.9952 % (> 99.9654 %)	96.2500 % (> 92.5000 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhe Bei Mu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Lian Qiao	5.12	–
Fu Zi	6.06	–
Tian Hua Fen	6.44	–
Ling Zhi	6.71	–
Pi Pa Ye	6.88	–
Bai Shao Yao	7.15	–
Ji Li	7.18	–
Shan Yao	7.37	–
Bai Zi Ren	7.93	–
Tai Zi Shen	9.64	–
He Huan Pi	10.04	–
Ze Xie	10.14	–
Jin Yin Hua	10.26	–
Di Gu Pi	11.03	–
Hou Po	11.45	–
Ren Dong Teng	11.89	–
Lai Fu Zi	12.02	–
Gua Lou	12.05	–
Ban Xia (Jiang)	12.16	–
Ye Jiao Teng	12.27	–
Gu Sui Bu	12.79	–
Huo Ma Ren	12.89	–
Mu Zei	12.95	–
Sheng Jiang	12.97	–
Shen Qu	13.00	–
Yan Hu Suo	14.00	–
Chuan Xiong	14.29	–
Tao Ren	14.32	–
Zhu Ru	14.33	–
Lian Zi	14.73	–
Ji Xue Teng	15.35	–
Jiao Gu Lan	15.41	–
Ku Shen	15.54	–
Gui Zhi	15.75	–
Yi Yi Ren	15.75	–
Fu Pen Zi	16.46	–
Che Qian Zi	16.49	–
Ci Wu Jia	16.53	–
Fu Xiao Mai	16.66	–
Cang Zhu	16.67	–
Tu Fu Ling	16.99	–
Ren Shen	17.33	–
San Qi	17.44	–
Yuan Zhi	17.46	–
Yu Jin	17.49	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Rou Gui	17.52	–
Chai Hu	17.52	–
Fu Ling	17.58	–
Wu Wei Zi	18.11	–
Suan Zao Ren	19.13	–
Ban Lan Gen	19.52	–
Zhu Ling	19.69	–
Ce Bai Ye	19.94	–
Dan Shen	20.02	–
Chen Pi	20.56	–
Lu Gen	20.60	–
Mao Dong Qing	20.71	–
Gou Teng	20.87	–
Dan Dou Chi	21.47	–
Dang Gui	22.44	–
Guang Huo Xiang	22.70	–
She Gan	22.81	–
Fo Shou	22.81	–
Bai Xian Pi	22.95	–
Zhi Ke	23.04	–
Qing Pi	23.51	–
Dang Gui Wei	23.96	–
Yu Zhu	24.12	–
Yin Yang Huo	24.49	–
Qiang Huo	24.73	–
Huang Qin	24.77	–
Gan Cao	24.83	–
Shan Yu Rou	25.13	–
Zhi Gan Cao	25.56	–
Xie Bai	25.76	–
Jie Geng	26.04	–
Gou Qi Zi	26.12	–
Ma Huang	26.45	–
Zhi Mu	26.87	–
(Fen) Bi Xie	28.78	–
Du Zhong	29.22	–
Hong Jing Tian	29.45	–
Zi Hua Di Ding	29.72	–
Chuang Mu Xiang	30.52	–
Huang Lian	31.12	–
Ma Huang Gen	32.22	–
Fu Shen	32.45	–
Mang Xiao	33.12	–
Mu Gua	34.03	–
Bo He	36.97	–
Sha Ren	37.47	–
Jing Jie	37.71	–
Chi Shao (Yao)	37.88	–
Sang Zhi	38.65	–
Cang Er Zi	39.46	–
Sang Ye	39.92	–
Long Yan Rou	41.39	–
Huang Bai	41.65	–
Hong Hua	43.36	–
Qing Hao	43.89	–
Bai Jiang Cao	48.41	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/substance group *Zhe Bei Mu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62687	62687	0.00	6.44
62688	62688	0.00	7.15
63027	63027	0.00	5.58
63028	63028	0.00	5.12

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhi Gan Cao**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60016-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhi Gan Cao; Glycyrrhizae radix et rhizoma praeparata cum melle

Special notes

When selecting the *Zhi Gan Cao* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhi Gan Cao	2	0	2

Second-stage model

For differentiation of the substance/substance group *Zhi Gan Cao* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhi Gan Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhi Gan Cao	G118HS089SV2	62869	40	from supplier
PhytoComm	Zhi Gan Cao	G118HS089SV2	62870	40	from supplier
PhytoComm	Zhi Gan Cao	G118H0881923	62899	40	from supplier
PhytoComm	Zhi Gan Cao	G118H0881923	62900	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Zhi Gan Cao*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.

- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Zhi Gan Cao*.

- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhi Gan Cao	G118HS089SV2	62869†	20
PhytoComm	Zhi Gan Cao	G118HS089SV2	62870†	20
PhytoComm	Zhi Gan Cao	G118H0881923	62899†	20
PhytoComm	Zhi Gan Cao	G118H0881923	62900†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 6 spectra from 4 *Apo-Ident* customers from 3 batches from the substance/substance group *Zhi Gan Cao*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Zhi Gan Cao	g118h0881322	1
PhytoComm	Zhi Gan Cao	G118H0881322	1
Phytocomm	Zhi Gan Cao	G118H0881522	4

- 851 spectra from 13 *Apo-Ident* customers from a total of 516 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhi Gan Cao* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhi Gan Cao* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	4	3	3	847

The substance/substance group *Zhi Gan Cao* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.7708 % (> 99.1837 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhi Gan Cao* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Huang Qi	6.59	–
Jin Yin Hua	7.15	–
Hong Jing Tian	7.38	–
Zi Hua Di Ding	9.60	–
Gan Cao	10.58	–
Jiao Gu Lan	11.16	–
Bai Zhu	11.29	–
Zhi Ke	11.68	–
Ban Lan Gen	11.83	–
Dan Shen	12.44	–
Dan Dou Chi	13.24	–
Hong Hua	13.82	–
Hou Po	14.18	–
Mu Gua	14.47	–
Jie Geng	14.56	–
Guang Huo Xiang	15.52	–
Ma Huang	15.57	–
Gou Qi Zi	15.87	–
Fu Zi	15.97	–
Chuan Xiong	15.97	–
Fu Pen Zi	16.01	–
Gua Lou	16.40	–
Qing Pi	16.75	–
Long Dan (Cao)	16.76	–
Chi Shao (Yao)	16.79	–
Chuang Mu Xiang	16.87	–
Shan Yao	17.16	–
Qiang Huo	17.67	–
Qin Jiao	17.67	–
(Huai) Niu Xi	17.70	–
Pi Pa Ye	17.74	–
Wu Wei Zi	17.95	–
Gu Sui Bu	18.20	–
Ren Dong Teng	18.27	–
Chai Hu	18.29	–
Suan Zao Ren	18.29	–
Bo He	18.32	–
Tian Hua Fen	18.38	–
Cang Zhu	18.90	–
Yan Hu Suo	18.92	–
Yin Yang Huo	19.16	–
Sang Zhi	19.22	–
Chuan Lian Zi	19.26	–
She Gan	19.32	–
Ye Jiao Teng	19.63	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	19.81	–
Huang Bai	19.82	–
Ren Shen	19.85	–
Mu Zei	20.22	–
Di Gu Pi	20.87	–
Bai Shao Yao	21.15	–
Lian Qiao	21.36	–
Sha Ren	21.65	–
Ku Shen	21.86	–
Ze Xie	22.28	–
Lu Gen	22.34	–
Mao Dong Qing	23.06	–
Chen Pi	23.33	–
Ji Li	23.99	–
Zhe Bei Mu	24.33	–
Huang Lian	24.77	–
Jing Jie	24.91	–
Lian Zi	24.97	–
Ce Bai Ye	25.16	–
Du Zhong	25.28	–
Shan Yu Rou	25.67	–
He Huan Pi	25.96	–
Qing Hao	26.20	–
Bai Xian Pi	26.25	–
Huang Qin	26.44	–
Xie Bai	26.45	–
Dang Gui	26.51	–
Sang Ye	27.45	–
Bai He	27.67	–
Yu Jin	27.93	–
Che Qian Zi	28.03	–
Shen Qu	28.11	–
Ji Xue Teng	28.47	–
Mang Xiao	28.58	–
Yu Zhu	28.64	–
Zi Su Zi	29.69	–
Bing Lang	30.05	–
Tao Ren	30.16	–
Long Yan Rou	30.49	–
Chuan Mu Tong	30.60	–
Cang Er Zi	30.95	–
San Qi	30.99	–
Bai Jiang Cao	31.12	–
Tai Zi Shen	31.20	–
Ling Zhi	31.21	–
Zhu Ling	31.65	–
Tu Fu Ling	31.76	–
Dang Gui Wei	31.78	–
(Fen) Bi Xie	31.86	–
Gou Teng	31.94	–
Chuan Niu Xi	32.41	–
Sheng Jiang	32.63	–
Gui Zhi	33.13	–
Bai Zhi	33.36	–
Bai Zi Ren	33.41	–
Rou Gui	33.48	–
Lai Fu Zi	33.59	–
Sha Shen (Bei)	34.17	–
(Shi) Chang Pu	35.07	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huo Ma Ren	35.98	–
Fu Ling	36.15	–
Gan Jiang	36.89	–
Zhi Mu	37.12	–
Ci Wu Jia	38.85	–
E Zhu	38.86	–
Pu Gong Ying	39.05	–
Ze Lan	41.12	–
Ban Zhi Lian	41.52	–
Ban Xia (Jiang)	43.43	–
Zhu Ru	43.86	–
Mi Huan Jun	43.96	–
Niu Bang Zi	44.48	–
Ma Huang Gen	44.77	–
Mai Ya	46.08	–
Ju Hua	47.82	–
Fang Feng	47.89	–
Yi Yi Ren	48.75	–
Wu Mei	48.83	–
Xiang Fu	49.54	–
Xi Xian Cao	50.00	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhi Gan Cao* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62869	62869	0.00	7.15
62870	62870	0.00	7.72
62899	62899	0.00	7.21
62900	62900	0.00	6.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhi Ke**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60158-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhi Ke; Citri aurantii fructus

Special notes

When selecting the *Zhi Ke* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhi Ke	2	0	2

Second-stage model

For differentiation of the substance/substance group *Zhi Ke* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhi Ke*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhi Ke	G043HS169RR1	62439	40	from supplier
PhytoComm	Zhi Ke	G043HS169RR1	62440	40	from supplier
PhytoComm	Zhi Ke	G043HS169SL1	62703	40	from supplier
PhytoComm	Zhi Ke	G043HS169SL1	62704	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Zhi Ke*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Zhi Ke*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhi Ke	G043HS169RR1	62439†	20
PhytoComm	Zhi Ke	G043HS169RR1	62440†	20
PhytoComm	Zhi Ke	G043HS169SL1	62703†	20
PhytoComm	Zhi Ke	G043HS169SL1	62704†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 3 spectra from 3 *Apo-Ident* customers from 2 batches from the substance/substance group *Zhi Ke*.
- Among them are spectra of independent samples from 2 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Zhi Ke	G043H0921221	1
Phytocomm	Zhi Ke	G043H0921422	2

- 854 spectra from 13 *Apo-Ident* customers from a total of 517 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhi Ke* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhi Ke* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	160	0	24 440
Type B	0	80	0	12 397
Type C	1	0	3	853

The substance/substance group *Zhi Ke* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9700 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9403 %)	100.0000 % (> 92.5000 %)
Type C	99.7674 % (> 99.1814 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhi Ke* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Jin Yin Hua	9.70	–
Gan Cao	12.22	–
Zhi Gan Cao	12.60	–
Lian Qiao	13.07	–
Jiao Gu Lan	13.27	–
Hou Po	13.83	–
Ma Huang	14.44	–
Zi Hua Di Ding	14.77	–
Ku Shen	15.89	–
Dan Dou Chi	17.09	–
Ban Lan Gen	17.57	–
Zhe Bei Mu	17.88	–
Cang Zhu	17.97	–
Yan Hu Suo	18.01	–
Fu Zi	18.43	–
Wu Wei Zi	18.59	–
Pi Pa Ye	18.63	–
San Qi	18.63	–
Fu Pen Zi	18.71	–
Guang Huo Xiang	18.95	–
Hong Jing Tian	19.22	–
Dan Shen	19.32	–
Yin Yang Huo	19.75	–
Shan Yao	19.83	–
Yu Zhu	19.88	–
Tian Hua Fen	19.99	–
Gou Qi Zi	20.10	–
Ye Jiao Teng	20.17	–
Suan Zao Ren	20.98	–
Qing Pi	21.26	–
Chuang Mu Xiang	21.64	–
Gua Lou	21.90	–
Dang Gui	22.21	–
Chai Hu	22.63	–
Tai Zi Shen	22.89	–
Shan Yu Rou	23.03	–
Qiang Huo	23.11	–
Lu Gen	23.74	–
Lai Fu Zi	24.24	–
Ren Dong Teng	24.26	–
Hong Hua	24.26	–
Bai Shao Yao	25.25	–
He Huan Pi	25.27	–
Huang Lian	25.37	–
Mu Zei	25.42	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chuan Xiong	25.48	–
She Gan	25.80	–
Huang Qin	25.81	–
Ling Zhi	25.82	–
Ze Xie	26.14	–
Bo He	26.35	–
Bai Zi Ren	27.07	–
Xie Bai	27.13	–
Huang Bai	27.26	–
Sha Ren	27.41	–
Ren Shen	27.79	–
Di Gu Pi	28.37	–
Chen Pi	28.59	–
Gu Sui Bu	28.96	–
Jing Jie	29.13	–
Ji Li	29.30	–
Lian Zi	29.34	–
Jie Geng	30.00	–
Sang Ye	30.15	–
Tao Ren	30.44	–
Mang Xiao	30.75	–
Qing Hao	31.26	–
Dang Gui Wei	31.30	–
Yuan Zhi	31.39	–
Che Qian Zi	31.69	–
Du Zhong	31.77	–
Bai Xian Pi	32.60	–
Ce Bai Ye	32.83	–
Mu Gua	32.94	–
Mao Dong Qing	33.15	–
Ji Xue Teng	33.82	–
Zhi Mu	34.40	–
Bai Jiang Cao	35.01	–
Rou Gui	35.15	–
Chi Shao (Yao)	35.21	–
Cang Er Zi	35.88	–
Yu Jin	36.73	–
Huo Ma Ren	37.63	–
Gui Zhi	37.68	–
Sheng Jiang	38.25	–
Zhu Ru	40.26	–
Ci Wu Jia	40.47	–
Zhu Ling	40.59	–
Shen Qu	40.97	–
(Fen) Bi Xie	41.06	–
Gou Teng	41.06	–
Yi Yi Ren	43.09	–
Fu Ling	43.42	–
Tu Fu Ling	43.57	–
Ze Lan	45.87	–
Long Yan Rou	47.86	–
Pu Gong Ying	48.36	–
Fu Xiao Mai	48.39	–
Wu Mei	48.91	–
Ban Xia (Jiang)	48.99	–
Ban Zhi Lian	49.40	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/substance group *Zhi Ke* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62439	62439	0.00	10.67
62440	62440	0.00	9.70
62703	62703	0.00	13.07
62704	62704	0.00	13.19

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhi Mu**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60196-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhi Mu; Anemarrhenae rhizoma

Special notes

When selecting the *Zhi Mu* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhi Mu	1	0	1

Second-stage model

For differentiation of the substance/substance group *Zhi Mu* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhi Mu*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhi Mu	G019HS153SW1	62959	40	from supplier
PhytoComm	Zhi Mu	G019HS153SW1	62960	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Zhi Mu*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Zhi Mu*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhi Mu	G019HS153SW1	62959†	20
PhytoComm	Zhi Mu	G019HS153SW1	62960†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 1 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Zhi Mu*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Zhi Mu	G019H0824321	1

- 856 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhi Mu* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhi Mu* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	77	3	24 520
Type B	0	36	4	12 437
Type C	0	0	1	856

The substance/substance group *Zhi Mu* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	96.2500 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	90.0000 % (> 82.5000 %)
Type C	100.0000 % (> 98.8367 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhi Mu* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yuan Zhi	2.72	–
Bai Shao Yao	7.11	–
Jie Geng	9.53	–
Pi Pa Ye	10.47	–
Shan Yao	10.90	–
Dang Gui Wei	12.78	–
San Qi	13.04	–
Fo Shou	13.33	–
Ren Shen	14.03	–
Gua Lou	14.39	–
Zhu Ru	14.55	–
Ye Jiao Teng	14.61	–
Ku Shen	14.74	–
Ji Li	14.99	–
Di Gu Pi	15.42	–
Mu Gua	15.67	–
Shen Qu	15.69	–
Fu Xiao Mai	16.47	–
Cang Zhu	16.93	–
Lian Zi	17.30	–
Dang Gui	18.50	–
He Huan Pi	18.62	–
Shan Yu Rou	19.14	–
Tian Hua Fen	19.64	–
Bai Zi Ren	19.69	–
Sheng Jiang	19.91	–
Rou Gui	19.94	–
Chuan Xiong	20.23	–
Ci Wu Jia	20.47	–
Lai Fu Zi	20.62	–
Lian Qiao	20.73	–
Fu Ling	21.43	–
Zhe Bei Mu	21.59	–
Ling Zhi	21.62	–
Gui Zhi	21.92	–
Ren Dong Teng	21.99	–
Gou Qi Zi	22.34	–
Tu Fu Ling	22.34	–
Mu Zei	22.53	–
Yu Zhu	22.74	–
Tai Zi Shen	23.09	–
Huo Ma Ren	23.39	–
Xie Bai	24.31	–
Chen Pi	25.01	–
Jiao Gu Lan	25.61	–
Yi Yi Ren	25.85	–
Ban Lan Gen	26.20	–
Suan Zao Ren	26.36	–
Jin Yin Hua	26.39	–
Tao Ren	26.40	–
Mai Men Dong	26.84	–
Mao Dong Qing	28.49	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Zi	29.34	–
Gou Teng	29.40	–
Ban Xia (Jiang)	29.53	–
Bai Xian Pi	29.63	–
Zhi Ke	30.13	–
Hou Po	30.51	–
Chuang Mu Xiang	30.74	–
Wu Wei Zi	30.81	–
Lu Gen	30.81	–
Ze Xie	30.83	–
Dan Dou Chi	31.28	–
Ji Xue Teng	31.56	–
Fu Pen Zi	31.93	–
Long Yan Rou	32.13	–
Dan Shen	33.22	–
Mang Xiao	33.54	–
Gu Sui Bu	33.57	–
Che Qian Zi	34.10	–
Huang Qin	34.56	–
Gan Cao	35.16	–
Yu Jin	36.30	–
Ma Huang Gen	38.21	–
Yan Hu Suo	38.92	–
Chai Hu	39.50	–
Hong Jing Tian	39.90	–
Ce Bai Ye	40.51	–
(Fen) Bi Xie	41.75	–
Yin Yang Huo	41.79	–
Guang Huo Xiang	42.66	–
Ma Huang	43.25	–
Fu Shen	43.50	–
Zhu Ling	43.84	–
Sang Zhi	44.43	–
Qiang Huo	45.88	–
Chi Shao (Yao)	46.88	–
She Gan	47.11	–
Qing Pi	50.08	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhi Mu* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62959	62959	0.00	2.72
62960	62960	0.00	3.34

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhi Shi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60145-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhi Shi; Aurantii fructus immaturus

Special notes

When selecting the *Zhi Shi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhi Shi	3	0	3

Second-stage model

For differentiation of the substance/substance group *Zhi Shi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhi Shi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhi Shi	G044HS168SK1	62655	40	from supplier
PhytoComm	Zhi Shi	G044HS168SK1	62656	40	from supplier
PhytoComm	Zhi Shi	G044HS168ST1	62979	40	from supplier
PhytoComm	Zhi Shi	G044HS168ST1	62980	40	from supplier
PhytoComm	Zhi Shi	G044HS168TK1	62981	40	from supplier
PhytoComm	Zhi Shi	G044HS168TK1	62982	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 240 spectra of 6 reference samples from the substance/substance group *Zhi Shi*. These samples are listed above in the *calibration samples* section. The reference samples come from 3 different batches.
- 24 360 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 120 spectra of 6 reference samples from the substance/substance group *Zhi Shi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhi Shi	G044HS168SK1	62655†	20
PhytoComm	Zhi Shi	G044HS168SK1	62656†	20
PhytoComm	Zhi Shi	G044HS168ST1	62979†	20
PhytoComm	Zhi Shi	G044HS168ST1	62980†	20
PhytoComm	Zhi Shi	G044HS168TK1	62981†	20
PhytoComm	Zhi Shi	G044HS168TK1	62982†	20

- 12 357 spectra from a total of 306 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 14 spectra from 6 *Apo-Ident* customers from 5 batches from the substance/substance group *Zhi Shi*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Phytocomm	Zhi Shi	g044h0922221	1
Phytocomm	Zhi Shi	G044H0922221	2
PhytoComm	Zhi Shi	G044H0922221	3
phytocomm	Zhi Shi	g044h0922321	1
Phytocomm	Zhi Shi	G044H0922321	2
PhytoComm	Zhi Shi	G044H0922321	2
Phytocomm	Zhi Shi	G044H0922521	3

- 843 spectra from 13 *Apo-Ident* customers from a total of 514 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhi Shi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhi Shi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	0	240	0	24 360
Type B	0	118	2	12 357
Type C	0	0	14	843

The substance/substance group *Zhi Shi* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9699 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9402 %)	98.3333 % (> 95.8333 %)
Type C	100.0000 % (> 98.8246 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhi Shi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yin Chen Hao	4.44	–
Sang Ye	6.02	–
Hu Zhang	8.78	–
Bai Hua She She Cao	8.88	–
Xin Yi	9.68	–
Xia Ku Cao	10.29	–
Ze Lan	12.69	–
Pu Gong Ying	13.61	–
Yu Xing Cao	15.48	–
Shan Yao	15.81	–
Nü Zhen Zi	16.12	–
Huang Lian	16.23	–
Han Lian Cao	16.26	–
Sha Ren	17.30	–
Bai Jiang Cao	17.72	–
Tu Fu Ling	21.58	–
Xi Xian Cao	21.76	–
Wu Mei	22.17	–
Qiang Huo	23.32	–
Huang Bai	24.57	–
Mang Xiao	24.76	–
Rou Cong Rong	25.55	–
Yi Mu Cao	26.30	–
Fu Zi	27.36	–
Jiang Huang	27.78	–
Du Zhong	28.13	–
Zi Hua Di Ding	28.20	–
Qing Pi	28.45	–
Bo He	29.34	–
Hong Jing Tian	29.52	–
Mao Dong Qing	29.81	–
Dan Shen	30.48	–
Qing Hao	30.52	–
Jing Jie	31.16	–
Ma Huang	31.20	–
Sang Ji Shend	31.42	–
Zhi Ke	32.08	–
Hong Hua	32.40	–
Gu Sui Bu	32.45	–
Wu Wei Zi	32.96	–
He Huan Pi	33.36	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Hou Po	33.42	–
Chai Hu	33.65	–
Dan Dou Chi	33.90	–
Gou Teng	34.77	–
Fu Pen Zi	36.02	–
Yin Yang Huo	36.87	–
Jiao Gu Lan	37.00	–
Cang Er Zi	37.29	–
Xiang Fu	38.18	–
Gan Cao	38.62	–
Che Qian Zi	40.17	–
Ye Jiao Teng	40.17	–
Yu Jin	40.74	–
Xuan Fu Hua	41.37	–
Chuang Mu Xiang	41.55	–
Niu Bang Zi	41.61	–
Zi Su Zi	41.71	–
Jin Yin Hua	42.05	–
Ce Bai Ye	42.38	–
Wang Bu Liu Xing	42.38	–
He Shou Wu	42.71	–
Wu Zhu Yu	43.69	–
Sang Bai Pi	44.04	–
Ju Hua	44.24	–
Yan Hu Suo	45.48	–
Shan Yu Rou	46.05	–
Gua Lou	46.50	–
Suan Zao Ren	46.52	–
Dan Zhu Ye	47.13	–
Bu Gu Zhi	47.57	–
Xuan Shen	47.88	–
Chuan Xiong	48.11	–
Lian Qiao	48.32	–
Zhe Bei Mu	48.32	–
Di Gu Pi	49.17	–
Xiao Hui Xiang	49.77	–
Zhi Gan Cao	50.06	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhi Shi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62655	62655	0.00	11.00
62656	62656	0.00	10.29
62979	62979	0.00	8.88
62980	62980	0.00	8.78

continued on the next page

continued from previous page

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62981	62981	0.00	5.50
62982	62982	0.00	4.44

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhu Ling**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	10002254-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhu Ling; Polypori umbellati sclerotium

Special notes

When selecting the *Zhu Ling* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhu Ling	1	0	3

Second-stage model

For differentiation of the substance/substance group *Zhu Ling* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhu Ling*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhu Ling	G199HS318TK1	62847	40	from supplier
PhytoComm	Zhu Ling	G199HS318TK1	62848	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Zhu Ling*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Zhu Ling*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhu Ling	G199HS318TK1	62847†	20
PhytoComm	Zhu Ling	G199HS318TK1	62848†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 5 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Zhu Ling*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
Mediherb	Zhu Ling		1
PhytoComm	Zhu Ling	G199H1509022	1
Phytocomm	Zhu Ling	G199H1509421	2
PhytoComm	Zhu Ling	G199H1509521	1

- 852 spectra from 12 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhu Ling* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhu Ling* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	0	0	5	852

The substance/substance group *Zhu Ling* can be clearly distinguished from all other substances. In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8263 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhu Ling* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Ling Zhi	6.86	–
Ren Dong Teng	7.03	–
Bai Xian Pi	7.35	–
Dan Dou Chi	8.49	–
Rou Gui	10.47	–
Gou Teng	10.81	–
She Gan	11.13	–
(Fen) Bi Xie	11.32	–
Huo Ma Ren	11.45	–
Ye Jiao Teng	11.52	–
Bai Shao Yao	11.96	–
Lian Zi	12.21	–
He Huan Pi	12.26	–
Ce Bai Ye	12.72	–
Lu Gen	13.11	–
Yi Yi Ren	13.29	–
Ma Huang Gen	13.42	–
Gui Zhi	13.63	–
Fu Ling	13.84	–
Sheng Jiang	14.29	–
Yu Jin	14.47	–
Guang Huo Xiang	14.66	–
Mao Dong Qing	14.68	–
Yin Yang Huo	15.00	–
Ji Xue Teng	15.35	–
Tu Fu Ling	15.43	–
Shen Qu	15.59	–
Mu Zei	16.25	–
Ji Li	16.55	–
Fu Xiao Mai	16.79	–
Chai Hu	16.87	–
Yan Hu Suo	17.11	–
Gu Sui Bu	17.69	–
Tai Zi Shen	18.06	–
Dan Shen	18.51	–
Chuan Xiong	18.93	–
Che Qian Zi	19.05	–
Tian Hua Fen	19.24	–
Tao Ren	19.57	–
Fu Zi	19.80	–
Lian Qiao	19.85	–
Fo Shou	20.01	–
Zhe Bei Mu	20.70	–
Hou Po	20.91	–
Pi Pa Ye	20.95	–
Shan Yao	21.00	–
Ci Wu Jia	21.19	–
Jing Jie	21.92	–
Ban Xia (Jiang)	22.60	–
Qiang Huo	22.86	–
Suan Zao Ren	23.30	–
Fu Pen Zi	24.23	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Di Gu Pi	24.28	–
Zi Hua Di Ding	24.75	–
Jiao Gu Lan	25.03	–
Gua Lou	25.14	–
Gan Cao	25.24	–
Ma Huang	25.28	–
Sha Ren	25.41	–
Bai Zi Ren	25.53	–
Fu Shen	26.50	–
Jin Yin Hua	26.59	–
Zhu Ru	27.02	–
Hong Jing Tian	28.25	–
Ban Lan Gen	29.57	–
Yuan Zhi	29.63	–
Du Zhong	29.63	–
Chen Pi	30.20	–
Lai Fu Zi	30.79	–
Zhi Ke	31.11	–
Cang Zhu	32.26	–
Huang Lian	32.84	–
Mang Xiao	32.87	–
Ze Xie	33.61	–
Ren Shen	33.70	–
Wu Wei Zi	33.79	–
Yu Zhu	33.81	–
Bo He	33.82	–
Qing Pi	35.20	–
Zhi Gan Cao	36.71	–
Dang Gui Wei	37.35	–
Huang Bai	37.53	–
Chuang Mu Xiang	38.26	–
Sang Zhi	39.97	–
Jie Geng	40.59	–
Dang Gui	41.11	–
Huang Qin	41.60	–
San Qi	41.99	–
Gou Qi Zi	43.55	–
Qing Hao	46.75	–
Long Yan Rou	47.06	–
Shan Yu Rou	47.19	–
Chi Shao (Yao)	47.24	–
Bai Jiang Cao	47.82	–
Ban Zhi Lian	49.73	–
Xie Bai	49.78	–
Bai Zhu	49.97	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhu Ling* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62847	62847	0.00	7.13
62848	62848	0.00	6.86

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zhu Ru**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60241-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zhu Ru; Bambusae in taeniae caulis

Special notes

When selecting the *Zhu Ru* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zhu Ru	2	0	3

Second-stage model

For differentiation of the substance/substance group *Zhu Ru* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zhu Ru*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhu Ru	G047HS109SK1	62657	40	from supplier
PhytoComm	Zhu Ru	G047HS109SK1	62658	40	from supplier
PhytoComm	Zhu Ru	G047HS109SW1	62913	40	from supplier
PhytoComm	Zhu Ru	G047HS109SW1	62914	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Zhu Ru*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Zhu Ru*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zhu Ru	G047HS109SK1	62657†	20
PhytoComm	Zhu Ru	G047HS109SK1	62658†	20
PhytoComm	Zhu Ru	G047HS109SW1	62913†	20
PhytoComm	Zhu Ru	G047HS109SW1	62914†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 9 spectra from 4 *Apo-Ident* customers from 4 batches from the substance/substance group *Zhu Ru*.
- Among them are spectra of independent samples from 3 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
phytocomm	Zhu Ru	g047h0625321	1
Phytocomm	Zhu Ru	g047h0625321	1
Phytocomm	Zhu Ru	G047H0625321	2
PhytoComm	Zhu Ru	G047H0625321	1
Phytocomm	Zhu Ru	G047H0625422	3
Phytocomm	Zhu Ru	G047h625321	1

- 848 spectra from 13 *Apo-Ident* customers from a total of 515 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zhu Ru* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zhu Ru* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type ($A/B/C$).

	False positive	True positive	False negative	True negative
Type A	3	154	6	24 437
Type B	3	58	22	12 394
Type C	0	0	9	848

The substance/substance group *Zhu Ru* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9929 % (> 99.9778 %)	96.2500 % (> 94.3750 %)
Type B	99.9810 % (> 99.9511 %)	72.5000 % (> 68.7500 %)
Type C	100.0000 % (> 98.8251 %)	n/a (n/a)

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zhu Ru* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Yi Yi Ren	3.09	–
Gui Zhi	3.89	–
Fu Xiao Mai	4.36	–
Sheng Jiang	4.85	–
Di Gu Pi	4.91	–
He Huan Pi	6.43	–
Ban Xia (Jiang)	6.93	–
Ci Wu Jia	7.53	–
Bai Shao Yao	7.81	–
Rou Gui	7.81	–
Fu Ling	8.10	–
Ji Li	8.97	–
Tao Ren	9.95	–
Shen Qu	10.22	–
Bai Zi Ren	11.07	–
Lian Zi	11.19	–
Gua Lou	12.57	–
Gou Teng	13.53	–
Lai Fu Zi	13.57	–
Zhe Bei Mu	13.77	–
Tai Zi Shen	13.93	–
Ling Zhi	14.16	–
Mu Zei	14.47	–
Tian Hua Fen	15.77	–
Fo Shou	15.94	–
Ye Jiao Teng	16.58	–
Yuan Zhi	16.81	–
Huo Ma Ren	16.85	–
Ji Xue Teng	16.89	–
Bai Xian Pi	16.94	–
Gu Sui Bu	17.69	–
Pi Pa Ye	18.94	–
Shan Yao	18.95	–
Ma Huang Gen	19.00	–
Lu Gen	20.41	–
Fu Shen	21.72	–
Yu Jin	22.18	–
Ren Dong Teng	23.12	–
Ze Xie	23.83	–
Tu Fu Ling	23.87	–
Lian Qiao	24.44	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chuan Xiong	25.15	–
Chen Pi	25.93	–
Zhu Ling	26.76	–
(Fen) Bi Xie	27.07	–
Zhi Mu	27.11	–
Jin Yin Hua	27.35	–
Suan Zao Ren	28.13	–
Dang Gui Wei	28.21	–
Yan Hu Suo	28.36	–
Ren Shen	28.38	–
Jie Geng	28.83	–
Shan Yu Rou	28.93	–
Fu Zi	30.72	–
Ce Bai Ye	31.08	–
San Qi	31.27	–
Cang Zhu	31.99	–
Fu Pen Zi	32.27	–
Yu Zhu	32.46	–
Jiao Gu Lan	32.65	–
Dang Gui	32.90	–
Dan Shen	33.20	–
Mang Xiao	33.62	–
Huang Qin	34.30	–
Dan Dou Chi	34.35	–
Mao Dong Qing	34.88	–
Ban Lan Gen	35.46	–
Zhi Ke	35.80	–
She Gan	36.12	–
Yin Yang Huo	36.25	–
Mu Gua	36.45	–
Chai Hu	37.78	–
Wu Wei Zi	38.79	–
Hou Po	38.85	–
Che Qian Zi	39.02	–
Ku Shen	39.64	–
Gou Qi Zi	40.44	–
Ma Huang	43.46	–
Hong Jing Tian	44.39	–
Qiang Huo	45.36	–
Long Yan Rou	46.06	–
Guang Huo Xiang	46.35	–
Gan Cao	46.46	–
Sang Zhi	47.47	–
Mai Men Dong	48.50	–
Huang Lian	48.72	–
Xie Bai	50.27	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhu Ru* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62657	62657	0.00	3.24
62658	62658	0.00	3.09
62913	62913	0.00	3.89
62914	62914	0.00	4.59

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zi Hua Di Ding**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	50376-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zi Hua Di Ding; Violae herba

Special notes

When selecting the *Zi Hua Di Ding* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zi Hua Di Ding	2	0	1

Second-stage model

For differentiation of the substance/substance group *Zi Hua Di Ding* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zi Hua Di Ding*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zi Hua Di Ding	G249HS269SK1	62627	40	from supplier
PhytoComm	Zi Hua Di Ding	G249HS269SK1	62628	40	from supplier
PhytoComm	Zi Hua Di Ding	G249HS269SW1	62803	40	from supplier
PhytoComm	Zi Hua Di Ding	G249HS269SW1	62804	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 160 spectra of 4 reference samples from the substance/substance group *Zi Hua Di Ding*. These samples are listed above in the *calibration samples* section. The reference samples come from 2 different batches.
- 24 440 spectra from a total of 307 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 80 spectra of 4 reference samples from the substance/substance group *Zi Hua Di Ding*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zi Hua Di Ding	G249HS269SK1	62627†	20
PhytoComm	Zi Hua Di Ding	G249HS269SK1	62628†	20
PhytoComm	Zi Hua Di Ding	G249HS269SW1	62803†	20
PhytoComm	Zi Hua Di Ding	G249HS269SW1	62804†	20

- 12 397 spectra from a total of 307 batches from further 175 substances. These spectra were

recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 2 spectra from 1 *Apo-Ident* customers from 1 batches from the substance/substance group *Zi Hua Di Ding*.
- Among them are spectra of independent samples from 1 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.

Supplier	Substance	Batch	Spectra
PhytoComm	Zi Hua Di Ding	G249H1216321	2

- 855 spectra from 13 *Apo-Ident* customers from a total of 518 batches from a further 215 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zi Hua Di Ding* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zi Hua Di Ding* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	1	160	0	24 439
Type B	1	78	2	12 396
Type C	0	0	2	855

The substance/substance group *Zi Hua Di Ding* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	99.9964 % (> 99.9814 %)	100.0000 % (> 96.2500 %)
Type B	99.9929 % (> 99.9630 %)	97.5000 % (> 93.7500 %)
Type C	100.0000 % (> 98.8302 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zi Hua Di Ding* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

Substance	Distance in main model	Distance in second-stage model
Fu Zi	4.84	–
Dan Shen	4.85	–
Qing Pi	7.19	–
Jiao Gu Lan	7.49	–
Qiang Huo	7.53	–
Hong Jing Tian	7.72	–
Chai Hu	7.92	–
Gan Cao	8.46	–
Fu Pen Zi	8.54	–
Bo He	8.59	–
Qing Hao	9.07	–
Hou Po	9.52	–
Jin Yin Hua	9.57	–
Mao Dong Qing	10.43	–
Jing Jie	10.50	–
Ce Bai Ye	11.27	–
Huang Bai	11.38	–
Dan Dou Chi	11.41	–
Du Zhong	12.44	–
Yin Yang Huo	12.70	–
Sha Ren	13.29	–
Yu Jin	13.34	–
Zhi Ke	13.39	–
Ban Lan Gen	13.93	–
Shan Yao	14.01	–
Zhi Gan Cao	14.17	–
Ma Huang	14.20	–
Guang Huo Xiang	14.24	–
Chuan Xiong	14.39	–
Ye Jiao Teng	14.56	–
Wu Wei Zi	14.99	–
Pi Pa Ye	15.31	–
Ling Zhi	15.87	–
Tian Hua Fen	15.99	–
Yan Hu Suo	16.53	–
She Gan	16.79	–
Ren Dong Teng	16.80	–
Gu Sui Bu	17.11	–
Bai Xian Pi	17.11	–
Ji Li	17.22	–
Bai Jiang Cao	17.50	–
Hong Hua	17.61	–
Huang Lian	17.86	–
Gua Lou	18.80	–
Shen Qu	19.11	–
Lian Zi	19.25	–
Bai Shao Yao	19.36	–
Zhe Bei Mu	19.59	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Suan Zao Ren	19.63	–
Che Qian Zi	19.91	–
Gou Qi Zi	19.93	–
Ren Shen	20.12	–
Di Gu Pi	20.44	–
Cang Zhu	20.86	–
Sang Ye	20.91	–
Lian Qiao	21.38	–
Ji Xue Teng	22.16	–
Mu Zei	22.53	–
Lu Gen	22.59	–
Chuang Mu Xiang	23.21	–
He Huan Pi	23.65	–
Yuan Zhi	23.71	–
(Fen) Bi Xie	24.26	–
Zhu Ling	24.41	–
Tu Fu Ling	25.13	–
Ze Xie	25.73	–
Shan Yu Rou	26.37	–
Cang Er Zi	26.71	–
Xie Bai	26.78	–
Jie Geng	26.81	–
Ku Shen	27.21	–
Chen Pi	27.22	–
Pu Gong Ying	27.38	–
Gou Teng	28.57	–
Chi Shao (Yao)	29.91	–
Sheng Jiang	30.21	–
Tao Ren	30.30	–
Mang Xiao	30.71	–
Ze Lan	30.73	–
Huang Qin	31.57	–
Fu Ling	31.92	–
Dang Gui	31.93	–
Huo Ma Ren	32.08	–
Xi Xian Cao	32.36	–
San Qi	34.19	–
Bai Zi Ren	34.39	–
Gui Zhi	35.07	–
Tai Zi Shen	35.20	–
Rou Gui	35.53	–
Sang Zhi	35.91	–
Wu Mei	36.54	–
Ma Huang Gen	36.58	–
Long Yan Rou	37.34	–
Yu Zhu	37.53	–
Nü Zhen Zi	38.99	–
Ban Zhi Lian	39.43	–
Ci Wu Jia	40.12	–
Dang Gui Wei	40.13	–
Ban Xia (Jiang)	41.34	–
Zhu Ru	42.16	–
Jiang Huang	43.00	–
Niu Bang Zi	43.84	–
Lai Fu Zi	44.05	–
Yi Yi Ren	44.13	–
E Zhu	44.41	–
Yi Mu Cao	44.90	–
Zhi Shi	45.16	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Mu	45.80	–
Zi Su Zi	46.84	–
Xin Yi	47.92	–
Mu Gua	48.52	–
(Shi) Chang Pu	48.89	–
Fu Xiao Mai	49.74	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zi Hua Di Ding* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62627	62627	0.00	7.19
62628	62628	0.00	4.84
62803	62803	0.00	7.49
62804	62804	0.00	7.53

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

VALIDATION REPORT

IdentModule 2.4-2022-01

Validated substance/substance group	**Zi Su Zi**
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	15/02/2022
Report number	60133-2022-02-15
Executing company	HiperScan GmbH Weißeritzstraße 3 01067 Dresden Germany

Relevant substance names

Zi Su Zi; Perillae fructus

Special notes

When selecting the *Zi Su Zi* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-7517-9	*European Pharmacopoeia 10th Edition, Basic Version 2020* [3]
Komm2.2.40	*Erfüllung von* Ph. Eur. 2.2.40 *durch* Apo-Ident [4]
AA004	*Erstellung und Validierung eines IdentModul-Updates*

Validation method

Validation is performed after every change to the chemometric model (also "database") in three steps:

1. The chemometric model is calculated from the calibration spectra using a *PCA algorithm*. The calibration spectra originate from the calibration samples of all substances in this class.

2. All spectra are presented to the generated chemometric model for evaluation. In three runs, the reference spectra (*Type A*), spectra from independent samples (*Type B*), and spectra from the field (*Type C*) are presented successively. Here, no single *false positive* result is permitted.

Finally, a report is generated from the validation run results and is archived together with the parameters of the model generation in audit-compliant manner.

Number of independent samples (batches) in calibration and validation

A sample is considered independent if no sample from the same batch has been included in the calibration of the chemometric model.

Substance/substance group	Type A	Type B	Type C
Zi Su Zi	1	0	0

Second-stage model

For differentiation of the substance/substance group *Zi Su Zi* the following second-stage model is used:

no second-stage model

Calibration samples

Only spectra which have been recorded by *HiperScan GmbH* from traceable samples are used for the generation of the chemometric models. The following samples have been obtained from the substance/substance group *Zi Su Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zi Su Zi	G183H1213922	62835	40	from supplier
PhytoComm	Zi Su Zi	G183H1213922	62836	40	from supplier

Validation samples

A total of 37 934 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 80 spectra of 2 reference samples from the substance/substance group *Zi Su Zi*. These samples are listed above in the *calibration samples* section. The reference samples come from 1 different batches.
- 24 520 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type A* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type A* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix A*. The samples in *Appendix A* were added for model generation not with the intention to have the model identify these substances, but with the intention to make the model reject them despite their similarity to substances identifiable.

Type B Spectra from independent samples not included in the database structure. Measurement is carried out by *HiperScan GmbH*.

Samples from batches of which no spectra have been included in the database structure are considered as independent samples. The number of batches from which independent samples supplied *Type B* spectra for validation is shown below, representing the number of independent *Type B* samples. Samples, of which some of the spectra have been included in the database structure and other spectra in the validation, are marked with a †. The following applies to the remaining unmarked samples: In the same batch, there was at least one additional sample (other sales container, other sample ID) from which reference spectra (*Type A*) were included in the database structure.

- 40 spectra of 2 reference samples from the substance/substance group *Zi Su Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. They are sorted upwards in the following table and separated from the additional samples by a line.

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Zi Su Zi	G183H1213922	62835†	20
PhytoComm	Zi Su Zi	G183H1213922	62836†	20

- 12 437 spectra from a total of 308 batches from further 175 substances. These spectra were recorded by *HiperScan GmbH*. *Type B* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type B* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix B*.

Type C Spectra of independent samples not included in the database structure. *Apo-Ident* customers carry out the measurements. *HiperScan GmbH* generally does not verify the information provided by the customer regarding manufacturer and batch number.

- 0 spectra from 0 *Apo-Ident* customers from 0 batches from the substance/substance group *Zi Su Zi*.
- Among them are spectra of independent samples from 0 batches from which no spectra have been included in the database structure. These are sorted upwards in the following table.
- 857 spectra from 13 *Apo-Ident* customers from a total of 519 batches from a further 216 substances. These spectra were recorded by *Apo-Ident* customers. *Type C* samples are documented within this entire report. If the substance can be identified with this chemometric model, the samples are listed in the *Type C* section of each substance page. If the substance cannot be identified by this model, the samples are listed in *Appendix C*.

Validation results

The validation runs checked whether the substance/substance group *Zi Su Zi* can be distinguished from all other substances using NIR spectroscopy with *Apo-Ident*. For this purpose, all relevant spectra of the various substances were compared with *Zi Su Zi* and it was evaluated how many matches (positive) and rejections (negative) were correct or incorrect. The following table breaks down the numbers of correct and incorrect results according to the expected result (*positive*/*negative*) and the validation spectrum type (*A*/*B*/*C*).

	False positive	True positive	False negative	True negative
Type A	0	80	0	24 520
Type B	0	40	0	12 437
Type C	1	0	0	856

The substance/substance group *Zi Su Zi* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable, the weighted *true negative rate* (*specificity*) and the weighted *true positive rate* (*recognition rate*) are determined:

	Specificity	Recognition rate
Type A	100.0000 % (> 99.9701 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9406 %)	100.0000 % (> 85.0000 %)
Type C	99.8843 % (> 99.3015 %)	n/a (n/a)

In order to ensure that each substance is entered with the same weight irrespective of the number of spectra available, each spectrum is weighted with the reciprocal value of its number. (If several new spectra of a substance with a very distinctive spectrum are added, the values for *specificity* and *recognition rate* are safer, but are by no means idealised.)

In order to illustrate the impact of the number of spectra, a comparison is made in parentheses as to how the *specificity* or *recognition rate* would worsen if three additional incorrect results were to exist amongst the spectra (*Rule of Three* [10, 11]). The larger the number of spectra, the lower the deterioration if the three hypothetical *incorrect results* were added.
Where the number of spectra does not exceed 20, no *recognition rate* is provided.

Nearest chemometric neighbours

The following table lists the closest chemometric neighbours of the substance/substance group *Zi Su Zi* in the substance class *TCM - Granulated herbal extracts (PhytoComm)*. Furthermore, their *Mahalanobis distance* is specified within the main model and, where appropriate, within the second-stage model.

[†]These spectra were recorded from material of the same package as the corresponding calibration spectra (*Type A*). Hence, these spectra are not independent and are not of *Type B*.

Substance	Distance in main model	Distance in second-stage model
Gan Jiang	5.89	–
Cang Er Zi	11.58	–
Sha Shen (Bei)	13.97	–
Ji Li	14.81	–
Chuan Lian Zi	15.64	–
Niu Bang Zi	15.82	–
Yan Hu Suo	17.61	–
Sang Zhi	17.81	–
(Huai) Niu Xi	18.57	–
Tian Hua Fen	18.70	–
Di Gu Pi	18.83	–
E Zhu	20.03	–
Jie Geng	21.56	–
(Shi) Chang Pu	22.28	–
Gua Lou	23.32	–
Jiang Huang	23.78	–
Shan Yao	24.44	–
Suan Zao Ren	24.82	–
Mi Huan Jun	26.27	–
Lai Fu Zi	26.61	–
Dang Gui	26.62	–
(Bai) Dou Kou	26.76	–
Lian Qiao	27.59	–
Xiao Hui Xiang	27.61	–
Mang Xiao	27.64	–
Xiang Fu	27.99	–
Chen Pi	28.02	–
Sha Ren	28.26	–
Huang Qi	29.66	–
Long Dan (Cao)	30.92	–
Chuan Mu Tong	30.93	–
Chuan Niu Xi	31.74	–
Ku Shen	32.31	–
Zhi Gan Cao	32.40	–
Bai Zhu	33.99	–
Mai Ya	35.24	–
Bai Zhi	35.33	–
Pu Gong Ying	36.97	–
Shan Yu Rou	37.45	–
Qin Jiao	37.88	–
Du Huo	37.92	–
Bai He	38.21	–
Bai Hua She She Cao	39.23	–
He Huan Pi	40.28	–
Dong Gua Zi	40.96	–
Yi Mu Cao	40.99	–
Wu Yao	41.70	–
Wu Wei Zi	42.52	–
Qiang Huo	42.63	–
Bai Xian Pi	42.64	–
Huang Lian	43.29	–
Jin Yin Hua	43.32	–
Fu Ling	43.84	–
Lian Zi	43.89	–
Tu Fu Ling	44.78	–
Zi Hua Di Ding	45.02	–
Ce Bai Ye	45.39	–
Ban Zhi Lian	45.63	–
Zhi Ke	45.82	–

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Da Zao	46.22	–
Chi Shao (Yao)	46.32	–
Bo He	46.67	–
Yuan Zhi	46.72	–
Fang Feng	47.73	–
Sang Ji Shend	48.18	–
Ba Ji Tian	48.33	–
Wang Bu Liu Xing	48.51	–
Mu Dan Pi	49.54	–
Xin Yi	49.78	–
Du Zhong	49.84	–
Ju Hua	50.03	–

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zi Su Zi* is separated from critical neighbours in a second-stage model, all the remaining substances in the second-stage model follow.

Identity according to identity test using NIR spectroscopy

For each substance, a certificate of the correct identity is obtained from an independent GMP-certified test laboratory. If the identity of the sample can be provided using NIR on an independently tested reference sample, in the following table the *Mahalanobis distance* to the reference sample is specified as well as the *Mahalanobis distance* to the next non-identical substance:

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62835	62835	0.00	5.89
62836	62836	0.00	6.48

The identity of a sample will be confirmed by NIR if the distance to the next foreign sample is at least 50 % greater than the distance to a reference sample whose identity has been established by laboratory testing. This criterion is always considered in the chemometric model, which contains all substances of the substance class, even if a second-stage dissolves a subgroup of similar substances, thereby increasing the distances between them. The samples confirmed by NIR support the statistical variance of the original reference substance, but cannot add new aspects or forms of the substance.

Appendix A: Additional calibration samples (*Type A*)

Not required.

Appendix B: Additional validation samples (*Type B*)

Not required.

Appendix C: Additional validation samples (*Type C*)

Spectra for substances from the field which cannot be identified with this model also enter the validation. In this manner, it is verified that the model also rejects unknown substances. The spectra for these samples were recorded by *Apo-Ident* customers. They belong to *Type C*. The information provided by the customer regarding the manufacturer and batch number is taken over by *HiperScan GmbH* to a large extent unchecked.

The samples originate from 116 batches. From these, 161 spectra were recorded. The spectra recorded on independent samples of substances from the field which can be identified with the model are listed respectively in the section *Type C* for the individual substances and do not appear again elsewhere in this list.

Supplier	Substance	Batch	Spectra
Phytocomm	(Bai) Jiang Can	g144h1708221	1
Phytocomm	(Bai) Jiang Can	g269h0513222	1
Phytocomm	(Bai) Jiang Can	G269H0513321	2
PhytoComm	(Bai) Jiang Can	G269H0513222	1

Supplier	Substance	Batch	Spectra
Phytocomm	Ai Ye	G033H0613022	1
Phytocomm	Ai Ye	G033H0613421	1
PhytoComm	Ai Ye	G033H0613021	1

Supplier	Substance	Batch	Spectra
Phytocomm	Bai Bian Dou	g093h0510221	1
Phytocomm	Bai Bian Dou	G093H0510221	1
Phytocomm	Bai Bu	G235H0608321	1
PhytoComm	Bai Bu	G235H0608221	1
Phytocomm	Bai Guo	G114H0509321	2
Phytocomm	Bai Guo	G114H0509521	1
Phytocomm	Bai Mao Gen	G125H0512221	2
Phytocomm	Bai Tou Weng	G207H0511521	4
PhytoComm	Bai Tou Weng	G207H0511221	3
PhytoComm	Bie Jia	G507H2401121	1
PhytoComm	Bie Jia	G507H2401321	1

Supplier	Substance	Batch	Spectra
PhytoComm	Cao Guo	E241050	1
phytocomm	Chan Tui	g067h1817322	1
Phytocomm	Chan Tui	G067H1817322	3
Phytocomm	Chan Tui	G067H1817323	2
Phytocomm	Chong Wei Zi	g317bp021046	1
Phytocomm	Chuan Bei Mu	g107f110721	1
Phytocomm	Chuan Bei Mu	G107BP020722A	1
Phytocomm	Chuan Bei Mu	g107f11021	1
PhytoComm	Chuan Bei Mu	G107F110721	1

Supplier	Substance	Batch	Spectra
Phytocomm	Da Fu Pi	G293H0321221	1
Phytocomm	Da Qing Ye	G127H0318421	2
EuRho	Dang Shen	g077h2001222	2
Phytocomm	Dang Shen	G077H2001022	1
Phytocomm	Dang Shen	G077H2001521	4
PhytoComm	Dang Shen	G077H2001621	1

continued on the next page

continued from previous page

Supplier	Substance	Batch	Spectra
Phytocomm	Deng Xin Cao	f100707	1
phytocomm	Di Long	g148h603321	1
Phytocomm	Di Long	G148H0603321	2
Phytocomm	Di Long	G148H0603421	2
PhytoComm	Di Long	G148H0603022	1
Phytocomm	Di Yu	G215H0604322	2

Supplier	Substance	Batch	Spectra
PhytoComm	E Bu Shi Cao	G061H180122	1
Phytocomm	E Jiao	g036h0861121	1
Phytocomm	E Jiao	g036h0861221	1
Phytocomm	E Jiao	G036H0861421	2
PhytoComm	E Jiao	G036H0861221	4

Supplier	Substance	Batch	Spectra
Phytocomm	Fang Ji	G236H0727321	1
PhytoComm	Fang Ji	G236H0727123	1

Supplier	Substance	Batch	Spectra
PhytoComm	Geng Mi	G329H1320121	1
Phytocomm	Gua Lou Ren	g243h1095123	1
PhytoComm	Gua Lou Ren	G243H1095421	1

Supplier	Substance	Batch	Spectra
PhytoComm	He Zi	G291H1239221	2
Phytocomm	Huai Hua	G263H1423421	1
Phytocomm	Huang Bo	h1203022	1
Phytocomm	Huang Bo	g188h1203123	1
PhytoComm	Huang Bo	G188H1203321	4
Phytocomm	Huang Jing	g196h1206121	1
Phytocomm	Huang Jing	G196H1206121	1

Supplier	Substance	Batch	Spectra
Phytocomm	Ji Nei Jin	G259H2102321	1
Phytocomm	Ji Nei Jin	g259h2102321	1
Phytocomm	Jie Cao	H2111921	1

Supplier	Substance	Batch	Spectra
PhytoComm	Kuan Dong Hua	H1219921	1

Supplier	Substance	Batch	Spectra
Phytocomm	Long Gu	g0251706122	1
PhytoComm	Long Gu	G025H1706122	1
Bios	Lu Hui	424263	1
phytocomm	Lu Jiao Jiao	G323H1158321	1
Phytocomm	Lu Jiao Shuang	FB1419901	1
Phytocomm	Lu Jiao Shuang	G062H1419421	1
Phytocomm	Lu Lu Tong	G142H1336321	1
Phytocomm	Lu Lu Tong	G142H1336521	2
PhytoComm	Lu Lu Tong	G142H1336122	1

Supplier	Substance	Batch	Spectra
Phytocomm	Ma Bian Cao	G248H1085221	1
PhytoComm	Ma Chi Xian	G255F110310	1
Phytocomm	Mo Yao	G1690732222	1
PhytoComm	Mo Yao	G169H0732222	2

Supplier	Substance	Batch	Spectra
Phytocomm	Qian Hu	G185H0926322	4
Phytocomm	Qian Shi	G102H0836422	1
PhytoComm	Qian Shi	G102H0836121	1
Phytocomm	Quan Xie	G222F110314	1

Supplier	Substance	Batch	Spectra
Phytocomm	Ru Xiang	G174H0820231	1
Phytocomm	Ru Xiang	G174H0820421	3

Supplier	Substance	Batch	Spectra
Phytocomm	San Leng	g270h0311221	1
PhytoComm	San Leng	G270H0311221	1
Phytocomm	Sang Piao Xiao	G282H421029201	1
Phytocomm	Sang Shen	G163H1047121	1
Phytocomm	Sang Shen	G163H1047221	2
Phytocomm	Shan Zha	G082H0331324	1
Phytocomm	Shan Zha	G082H0331521	1
Phytocomm	She Chuang Zi	G076H1154122	1
PhytoComm	She Chuang Zi	G076H1154021	1
Phytocomm	Sheng Ma	g068h04317121	1
Phytocomm	Sheng Ma	G068H0437421	2
PhytoComm	Suo Yang	G260H1831121	1

Supplier	Substance	Batch	Spectra
Phytocomm	Tian Men Dong	G037H0415221	1
Phytocomm	Tian Men Dong	g037h0415221	1
Phytocomm	Tian Men Dong	g03h0415521	1
Phytocomm	Tian Men Dong	G037H0415321	1

Supplier	Substance	Batch	Spectra
Phytocomm	Wei Ling Xian	G075H0943223	3
Phytocomm	Wei Ling Xian	g075h0943223	1
Phytocomm	Wei Ling Xian	G075H0943422	1
PhytoComm	Wei Ling Xian	GO75H0943223	1
Phytocomm	Wu Gong	G311H1411321	1
Phytocomm	Wu Gong	G311H411321	1

Supplier	Substance	Batch	Spectra
Phytocomm	Xi Yang Shen	G324H0647321	1
Phytocomm	Xi Yang Shen	g324h0647321	1
Phytocomm	Xi Yang Shen	G324HS407OL1	1
phytocomm	Xian He Cao	g008h0547321	1
Phytocomm	Xian He Cao	G008H0547121	1
Phytocomm	Xian He Cao	G008H0547321	1

Supplier	Substance	Batch	Spectra
Phytocomm	Yi Zhi Ren	G015H1051321	2
Phytocomm	Yi Zhi Ren	G015H1051521	1

Supplier	Substance	Batch	Spectra
Phytocomm	Zhen Zhu Mu	g155h0979221	1
Phytocomm	Zhi Zi	g110h1137121	1
Phytocomm	Zhi Zi	G110H1137521	4
PhytoComm	Zhi Zi	G110H1137121	1
PhytoComm	Zi Su Ye	G182H1214321	3
PhytoComm	Zi Su Ye	G182HS268OP1	1
Phytocomm	Zi Wan	G038H1215321	2

Appendix D: Requirements of validation

In order to ensure adherence to the safe scientific status, the individual methods for manufacturing and testing must be validated under certain circumstances (compare § 34 para. 1 no. 3, § 35 para. 1 no. 4 and para. 4 sentence 1 no. 2 b, para. 6 sentence 3 *ApoBetrO* [Pharmacies Rules and Regulations]). The *ApoBetrO* [Pharmacies Rules and Regulations] incorporates a legal definition in § 1 a para. 16 (quotation translated):

> *"Validation is the provision of documented proof which with a high degree of safety documents that, via a specific process or standard work process, a medicinal product is manufactured and tested, which is in accordance with previously determined quality features."*

Validation documentation can be used to prove that methods or devices which are not described in the Pharmacopoeia within the meaning of § 6 para. 1 sentence 3 *ApBetrO* [Pharmacies Rules and Regulations] achieve the same results as those in the Pharmacopoeia. On the other hand, with the requirements of the demanded validation it must be observed whether the respective testing method is already incorporated in the Pharmacopoeia.

NIR spectroscopy as a general testing method need not be validated in accordance with the express ruling in the *Ph. Eur. Section 1.1* [3], as it is already described in *Section 2.2.40* of the *Ph. Eur.* as an area of application for the identification of raw materials.

However, a special validation requirement exists for the reference database. This requirement is met with the existing document. Further requirements or rules as to how this proof must be furnished do not exist. It is required that the processes guarantee the same results as the methods and devices in the Pharmacopoeia [17].

Carrying out identity tests with *Apo-Ident* is therefore also possible if the NIR spectroscopy process is not required in the Pharmacopoeia monograph of the substance for identity testing. All NIR analyses with *Apo-Ident* prove several, often all molecule groups and are therefore comparable with a series of individual, targeted chemical proofs [4]. Therefore, the identity proof with *Apo-Ident* replaces the monograph test series (with two or more test combinations).

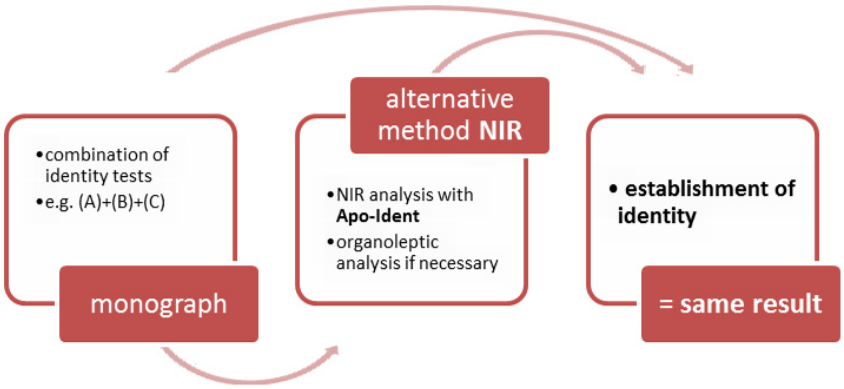

Figure 2: The combination of tests of the monograph is replaced by the alternative method NIR spectroscopy using *Apo-Ident*. This is permissible because both test procedures result in the establishment of the identity of the raw material.

With this validation documentation, proof is furnished that identical results are achieved with *Apo-Ident* and Pharmacopoeia methods, i.e. confirmation of the identity of the raw material [2].

Appendix E: Conformity of *Apo-Ident* with the *European pharmacopeia*

According to *Ph. Eur. Section 2.2.40*, NIR spectroscopy is basically suitable for: "Identification of agents, excipients, dosage forms, intermediate manufacturing products, chemical raw materials and packaging materials" ([3], quotation translated).

The fact that *Apo-Ident* meets the further criteria of the European Pharmacopoeia under the headings in *Section 2.2.40*

- Apparatus
- Measurement methods
- Sample preparation and presentation
- Testing the functionality of the instrument
- Identification and characterisation (qualitative analysis)
- Quantitative analysis
- Ongoing model evaluation
- Transfer of databases
- Data storage

can be proven based on the *HiperScan GmbH* documentation of "Meeting *2.2.40 Ph. Eur.* by *Apo-Ident*" [4].

References

[1] ABDA – Bundesvereinigung Deutscher Apothekerverbände: Verordnung über den Betrieb von Apotheken (Apothekenbetriebsordnung – ApBetrO), 2012

[2] Reimann, B. ; Regierungspräsidium Darmstadt: Hinweise zur ordnungsgemäßen Prüfung von Arzneimitteln und Ausgangsstoffen (§§ 6 und 11 *ApBetrO*), 2007

[3] *Europäisches Arzneibuch, Grundwerk 2020*. 10. Ausgabe. Deutscher Apotheker Verlag (978-3-7692-7517-9)

[4] HiperScan GmbH: Erfüllung von 2.2.40 Ph. Eur. durch Apo-Ident, 2013

[5] Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD): Resolution in Verbindung mit §§ 6 und 11 *ApBetrO* Verwendung eines Nah-Infrarot-Spektrometers (NIR) zur Identitätsprüfung, 16. 10. 2013, DAZ 48, November 2013

[6] Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD): Resolution 2014, Arbeitsgemeinschaft der Pharmazieräte Deutschlands (APD), Oktober 2014

[7] *DAC/NRF*. Govi-Verlag (978-3-7741-0044-2)

[8] Kessler, W.: *Multivariate Datenanalyse*. WILEY-VCH Verlag, 2007 (978-3-527-31262-7)

[9] Næs, T. ; Isaksson, T. ; Fearn, T. ; Davies, T.: *Multivariate Calibration and Classification*. NIR Publications, 2002 (978 0 9528666 2 6)

[10] Hanley, J. A. ; Lippman-Hand, A.: If nothing goes wrong, is everything all right? In: *Journal of the American Medical Association* 249 (1983), S. 1743–1745

[11] Jovanovic, B. D. ; Levy, P. S.: A Look at the Rule of Three. In: *The American Statistician* 51 (1997), S. 137–139

[12] Bronstein, I. N. ; Semendjajew, K. A. ; Musiol, G. ; Mühlig, H.: *Taschenbuch der Mathematik*. 5. überarbeitete und erweiterte Auflage. Verlag Harri Deutsch, 2000 (3-8171-2015-2)

[13] Mahalanobis, P.: On the generalized distance in statistics. In: *Proc. Nat. Inst. Sci. India (Calcutta)* 2 (1936), S. 49–55

[14] Yambor, B. ; Draper, W. ; Beveridge, R.: Analyzing PCA-based face recognition algorithms: Eigenvector selection and distance measures. In: *Second Workshop Empirical Evaluation in Computer Vision* (2000)

[15] HiperScan GmbH: Identifikationsmethodik Apo-Ident, 2012

[16] HiperScan GmbH: Datenvorbehandlung des Identifikationssystems Apo-Ident, 2012

[17] Cyran ; Rotta: Apothekenbetriebsordnung, Kommentar § 6 Rn. 10, 2010

Index