

Validation documentation
TCM - Granulated herbal extracts (PhytoComm)

HiperScan GmbH

March 12, 2020

Contents

Introduction	7
Context of this document	7
Criteria for the inclusion of substances	8
Validation concept	9
Model creation procedure and validation runs	9
Particularities of individual substance classes	12
Significance of testing with <i>Apo-Ident</i>	13
Conclusion	13
Explanation of terminology	14
Summary	19
Validation samples	19
Validation results	19
Validation reports	21
(Bai) Dou Kou	21
(Huai) Niu Xi	27
(Ku) Xing Ren	33
(Sheng) Di Huang	37
(Shi) Chang Pu	43
Ai Ye	49
Ba Ji Tian	53
Bai Bian Dou	59
Bai He	65
Bai Hua She She Cao	71
Bai Mao Gen	77
Bai Shao Yao	83
Bai Xian Pi	89
Bai Zhu	95
Bai Zi Ren	101
Ban Xia (Jiang)	107
Ban Zhi Lian	113
Bu Gu Zhi	119
Cang Er Zi	125
Cang Zhu	131
Chai Hu	137
Che Qian Zi	145
Chen Pi	151
Chi Shao (Yao)	159
Chuan Lian Zi	165
Chuan Xiong	171
Chuang Mu Xiang	179
Ci Wu Jia	185
Da Huang	191
Da Zao	195
Dan Dou Chi	201
Dan Shen	207
Dan Zhu Ye	215
Dang Gui	221
Dang Gui Wei	227
Dang Shen	233
Di Fu Zi	239
Di Gu Pi	243
Dong Gua Zi	249
Dong Kui Zi	253

Du Huo	257
Du Zhong	263
E Zhu	269
Fang Feng	275
Fang Ji	281
Fo Shou	287
Fu Ling	293
Fu Shen	299
Fu Xiao Mai	305
Fu Zi	311
Gan Cao	319
Gan Jiang	325
Ge Gen	331
Gou Qi Zi	337
Gou Teng	343
Gu Sui Bu	349
Gua Lou	355
Guang Huo Xiang	363
Gui Zhi	367
Han Lian Cao	375
He Huan Pi	381
He Shou Wu	387
Hong Hua	393
Hong Jing Tian	399
Hou Po	405
Hu Zhang	411
Hua Shi	417
Huang Bai	421
Huang Bo	427
Huang Jing	433
Huang Lian	439
Huang Qi	445
Huang Qin	451
Huo Ma Ren	457
Ji Li	463
Ji Xue Teng	471
Jiang Huang	477
Jiao Gu Lan	483
Jie Geng	489
Jin Qian Cao	495
Jin Yin Hua	501
Jing Jie	507
Ju Hua	513
Jue Ming Zi	519
Ku Shen	523
Lai Fu Zi	529
Lian Qiao	535
Lian Zi	541
Ling Zhi	547
Long Dan (Cao)	553
Long Yan Rou	559
Lu Gen	565
Ma Bian Cao	571
Ma Huang	577
Ma Huang Gen	583
Mai Men Dong	589
Mai Ya	595

Man Jing Zi	601
Mang Xiao	605
Mao Dong Qing	609
Mo Yao	615
Mu Dan Pi	621
Mu Gua	627
Niu Bang Zi	633
Nü Zhen Zi	639
Pi Pa Ye	645
Pu Gong Ying	651
Qiang Huo	657
Qin Jiao	663
Qing Hao	669
Qing Pi	673
Ren Dong Teng	679
Rou Cong Rong	685
Rou Gui	691
Ru Xiang	697
San Leng	701
San Qi	707
Sang Bai Pi	713
Sang Ji Shend	719
Sang Ye	725
Sang Zhi	731
Sha Ren	737
Sha Shen (Bei)	743
Shan Yao	749
Shan Yu Rou	755
She Gan	761
Shen Qu	767
Sheng Jiang	773
Shi Gao	779
Shu Di (Huang)	783
Suan Zao Ren	787
Tai Zi Shen	793
Tao Ren	799
Tian Hua Fen	805
Tian Men Dong	811
Tu Fu Ling	817
Tu Si Zi	825
Wang Bu Liu Xing	829
Wei Ling Xian	835
Wu Jia Pi	841
Wu Mei	845
Wu Wei Zi	851
Wu Yao	857
Wu Zhu Yu	863
Xi Xian Cao	869
Xia Ku Cao	875
Xian He Cao	881
Xiang Fu	887
Xie Bai	893
Xin Yi	899
Xu Duan	905
Xuan Shen	911
Yan Hu Suo	917
Ye Jiao Teng	925

Yi Mu Cao	931
Yi Yi Ren	937
Yin Chen Hao	943
Yin Yang Huo	949
Yu Jin	955
Yu Xing Cao	961
Yu Zhu	965
Yuan Zhi	971
Ze Xie	977
Zhe Bei Mu	983
Zhi Ke	989
Zhi Mu	995
Zhi Shi	1001
Zhi Zi	1007
Zhu Ling	1013
Zhu Ru	1019
Zi Hua Di Ding	1025
Zi Su Zi	1031
Zi Wan	1037
Appendix	1043
Additional calibration samples (<i>Type A</i>)	1043
Additional validation samples (<i>Type B</i>)	1045
Additional validation samples (<i>Type C</i>)	1047
Requirements of validation	1051
Conformity of <i>Apo-Ident</i> with the <i>European pharmacopeia</i>	1052
References	1053
Index	1055

Introduction

The unequivocal proof of the identity of pharmaceutical raw materials based on a monograph or traditional alternative methods is work-intensive, time-intensive 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 transfection.
- 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. Bezugsquellen-nachweis 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) *Unterschnittener 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 themselves 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 Pharmacopoeia* 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

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

[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 $\mathbf{X}$.

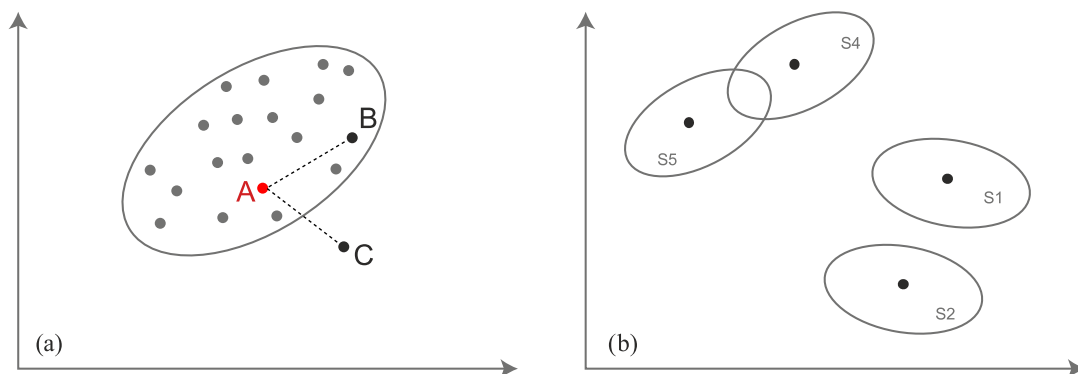


Figure 1: (a) The *Mahalanobis distance* from *A* to *B* is less than from *A* to *C*. However, the *Euclidean distances* are identical. (b) The *Mahalanobis distance* between the two measurement series *S4* and *S5* is smaller than between *S1* and *S2*. However, the *Euclidean distances* are identical.

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 39 818 spectra from 830 different batches for a total of 233 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)	173	319	25 801

From category *A* a total of 25 801 spectra from 319 batches for a total of 173 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)	173	320	13 167

From category *B* a total of 13 167 spectra from 320 batches for a total of 173 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)	215	514	850

From category *C* a total of 850 spectra from 514 batches for a total of 215 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	101	25 609	192	4 437 671
Type B	192	12 832	335	2 264 532
Type C	187	107	598	146 677

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	99.997 90 % (> 99.997 81 %)	99.310 14 % (> 99.295 32 %)
Typ B	99.991 96 % (> 99.991 79 %)	97.827 01 % (> 97.798 31 %)
Typ C	99.864 10 % (> 99.860 71 %)	12.238 86 % (> 11.560 37 %)

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 2.1-2020-01

Validated substance/substance group **(Bai) Dou Kou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002257-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 (*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	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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
Gui Zhi	12.12	—
Jiang Huang	15.34	—
Tu Fu Ling	16.30	—
Sha Ren	16.52	—
Ji Li	18.14	—
He Huan Pi	21.22	—
Mang Xiao	21.67	—
E Zhu	23.24	—
Sha Shen (Bei)	23.53	—
Yu Jin	24.31	—
Sang Zhi	25.97	—
Chuan Xiong	27.67	—
Fu Zi	28.30	—
Yi Mu Cao	28.35	—
Wu Yao	28.97	—
Zhu Ling	29.89	—
Xian He Cao	30.83	—
Cang Er Zi	30.86	—
Ba Ji Tian	31.07	—
Shan Yao	31.17	—
Chen Pi	31.42	—
Yin Yang Huo	32.45	—
Long Dan (Cao)	32.91	—
Yan Hu Suo	33.66	—
Mu Dan Pi	34.74	—
Sheng Jiang	34.91	—
Huang Lian	36.16	—
Lu Gen	36.24	—
Chai Hu	36.29	—
Zi Su Zi	36.80	—
He Shou Wu	37.23	—
Gan Jiang	37.29	—
She Gan	38.91	—
Bu Gu Zhi	39.07	—
Xiang Fu	39.15	—
Chuan Lian Zi	39.24	—
Ma Bian Cao	39.85	—
Di Gu Pi	40.58	—
Du Huo	40.94	—
Zhi Ke	41.15	—
Xia Ku Cao	41.15	—
Xin Yi	41.45	—
Bai Bian Dou	41.95	—
(Huai) Niu Xi	42.12	—
Niu Bang Zi	42.20	—
Hu Zhang	42.32	—
Yuan Zhi	42.35	—
Bai Xian Pi	42.54	—
Sang Ji Shend	42.81	—
Mai Ya	42.99	—
Wang Bu Liu Xing	43.12	—
Sang Bai Pi	43.14	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mo Yao	43.53	—
Xi Xian Cao	43.55	—
Huo Ma Ren	44.14	—
San Qi	44.22	—
Hou Po	44.27	—
Gua Lou	44.36	—
San Leng	44.51	—
Fang Feng	44.57	—
Fang Ji	44.86	—
Lai Fu Zi	44.86	—
Bai He	44.92	—
Xu Duan	44.97	—
Dang Gui	45.42	—
Qiang Huo	45.53	—
Dan Shen	45.60	—
(Shi) Chang Pu	45.67	—
Fu Ling	45.91	—
Du Zhong	45.96	—
Qing Pi	46.84	—
Zi Wan	46.91	—
Dang Gui Wei	46.91	—
Hong Jing Tian	46.92	—
Xie Bai	46.95	—
Lian Qiao	47.29	—
Gu Sui Bu	48.06	—
Yin Chen Hao	48.58	—
Huang Jing	49.00	—
Huang Bo	49.12	—
Gou Qi Zi	49.26	—
Dan Dou Chi	49.85	—

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
62762	62762	0.00	12.12
62761	62761	0.00	12.12

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 2.1-2020-01

Validated substance/substance group (Huai) Niu Xi
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002255-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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	G003H1905222	1
Phytocomm	(Huai) Niu Xi	g003h1905222	1
Phytocomm	(Huai) Niu Xi	G003H1905321	3
Phytocomm	(Huai) Niu Xi	G003H1905422	1
PhytoComm	(Huai) Niu Xi	G003H1905222	1
PhytoComm	(Huai) Niu Xi	G003H1905222	1
PhytoComm	(Huai) Niu Xi	G003H1905422	1

- 844 spectra from 12 *Apo-Ident* customers from a total of 511 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	25 721
Type B	0	40	0	13 127
Type C	5	0	9	839

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.5039 % (> 98.9153 %)	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	11.17	—
Gua Lou	13.27	—
Zhi Mu	13.70	—
Huang Qi	13.84	—
She Gan	14.14	—
Bai Mao Gen	14.55	—
Dang Gui Wei	15.27	—
(Shi) Chang Pu	15.31	—
Zi Wan	15.33	—
Jie Geng	15.75	—
Sang Zhi	16.57	—
Mu Gua	18.54	—
Gou Qi Zi	18.85	—
Qing Pi	18.93	—
Cang Zhu	19.03	—
Di Gu Pi	19.82	—
Dang Shen	20.08	—
Chai Hu	20.34	—
Lian Qiao	20.66	—
San Leng	22.26	—
Bai Zhu	24.30	—
Gan Jiang	24.70	—
Chi Shao (Yao)	25.00	—
Ku Shen	25.16	—
Niu Bang Zi	26.04	—
Gui Zhi	26.10	—
Xiang Fu	26.35	—
Chuan Xiong	27.29	—
Huo Ma Ren	27.49	—
Du Huo	28.11	—
Ze Xie	28.46	—
Yi Mu Cao	29.05	—
Tian Hua Fen	29.05	—
Fang Feng	29.66	—
Xie Bai	29.69	—
Lai Fu Zi	29.77	—
Dang Gui	30.43	—
Mang Xiao	30.72	—
Ju Hua	31.00	—
Lu Gen	31.74	—
Hong Jing Tian	32.11	—
Wei Ling Xian	32.33	—
Ji Li	32.36	—
Mu Dan Pi	32.42	—
Sheng Jiang	33.03	—
Zhi Ke	33.82	—
Zi Su Zi	34.64	—
Bai Shao Yao	34.68	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yao	35.93	—
Qin Jiao	35.99	—
Hou Po	36.13	—
Long Dan (Cao)	36.19	—
Ba Ji Tian	36.44	—
Bai He	36.86	—
Cang Er Zi	37.20	—
Wu Yao	37.28	—
San Qi	37.38	—
E Zhu	37.56	—
Jiang Huang	38.01	—
Zhi Zi	38.45	—
Jin Yin Hua	38.48	—
Dong Gua Zi	38.60	—
Chen Pi	38.79	—
Gan Cao	39.77	—
Lian Zi	40.34	—
Shan Yu Rou	41.13	—
Bai Bian Dou	41.77	—
Wu Wei Zi	41.83	—
Sha Shen (Bei)	42.04	—
Chuan Lian Zi	42.76	—
Bai Xian Pi	42.95	—
Fu Zi	43.65	—
Yan Hu Suo	44.43	—
Zhe Bei Mu	44.62	—
Pu Gong Ying	46.40	—
(Bai) Dou Kou	46.44	—
Xian He Cao	46.70	—
Huang Bo	47.04	—
Sha Ren	48.02	—
Mai Men Dong	48.33	—
Qiang Huo	48.70	—
Tu Fu Ling	48.94	—
Huang Lian	48.96	—
Xin Yi	49.46	—

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
62760	62760	0.00	12.92
62759	62759	0.00	11.17

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 2.1-2020-01

Validated substance/substance group (Ku) Xing Ren
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50290-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G029H0708821	62393	40	from supplier
PhytoComm	(Ku) Xing Ren	G029H0708821	62394	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G029H0708821	62393 [†]	20
PhytoComm	(Ku) Xing Ren	G029H0708821	62394 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	g029h0708322	2
PhytoComm	(Ku) Xing Ren	G029H0708125	2
PhytoComm	(Ku) Xing Ren	g029h0708122	1
PhytoComm	(Ku) Xing Ren	G029H0708125	1
PhytoComm	(Ku) Xing Ren	G029H0708125	1
PhytoComm	(Ku) Xing Ren	G029H0708322	2
PhytoComm	(Ku) Xing Ren	G029H0708622	1

- 843 spectra from 12 *Apo-Ident* customers from a total of 510 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	25 561
Type B	0	120	0	13 047
Type C	0	0	10	843

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8227 %)	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
Ru Xiang	52.81	—

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
62510	62510	0.00	61.06
62767	62767	0.00	59.38
62768	62768	0.00	59.43
62509	62509	0.00	61.03
62394	62394	0.00	53.66
62393	62393	0.00	52.81

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 2.1-2020-01

Validated substance/substance group (Sheng) Di Huang
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60005-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
(Sheng) Di Huang	3	0	4

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	G211H0532621	61797	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532621	61804	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532621	61881	40	from supplier
PhytoComm	(Sheng) Di Huang	G211H0532621	61882	40	from supplier
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

Validation samples

A total of 39 821 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 *(Sheng) Di Huang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 481 spectra from a total of 316 batches from further 172 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 *(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	G211H0532621	61797 [†]	20
PhytoComm	(Sheng) Di Huang	G211H0532621	61804 [†]	20
PhytoComm	(Sheng) Di Huang	G211H0532621	61881 [†]	20
PhytoComm	(Sheng) Di Huang	G211H0532621	61882 [†]	20

continued on the next page

continued from previous page

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

- 13 007 spectra from a total of 317 batches from further 172 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 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	(Sheng) Di Huang	g211h053221	1
Phytocomm	(Sheng) Di Huang	g211h0532121	1
Phytocomm	(Sheng) Di Huang	G211H0532521	4
	(Sheng) Di Huang	g211h0532323	1
PhytoComm	(Sheng) Di Huang	G211H0532621	1

- 845 spectra from 12 *Apo-Ident* customers from a total of 511 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	320	0	25 481
Type B	0	160	0	13 007
Type C	0	2	6	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 (*Sheng*) *Di 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 % (> 98.1250 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 96.2500 %)
Type C	100.0000 % (> 98.8230 %)	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 (*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	8.97	—
Wu Jia Pi	19.02	—
Pu Gong Ying	24.15	—
Shu Di (Huang)	25.34	—
Sha Ren	27.73	—
Ai Ye	29.54	—
Dan Zhu Ye	29.85	—
Bu Gu Zhi	34.70	—
Xuan Shen	35.63	—
Sang Bai Pi	35.80	—
Yin Chen Hao	37.92	—
Jing Jie	39.10	—
Guang Huo Xiang	39.24	—
Jue Ming Zi	43.11	—
Hua Shi	44.88	—
Ge Gen	46.57	—
Gu Sui Bu	51.50	—

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
62750	62750	0.00	11.43
62312	62312	0.00	9.16
61804	61804	0.00	13.62
62311	62311	0.00	8.97
62749	62749	0.00	11.21
61882	61882	0.00	15.42
61797	61797	0.00	12.00
61881	61881	0.00	15.51

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 2.1-2020-01

Validated substance/substance group (Shi) Chang Pu
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60426-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G006H0544622	62107	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544622	62108	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544821	62399	40	from supplier
PhytoComm	(Shi) Chang Pu	G006H0544821	62400	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G006H0544622	62107 [†]	20
PhytoComm	(Shi) Chang Pu	G006H0544622	62108 [†]	20
PhytoComm	(Shi) Chang Pu	G006H0544821	62399 [†]	20
PhytoComm	(Shi) Chang Pu	G006H0544821	62400 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G006H0544122	1
Phytocomm	(Shi) Chang Pu	G006H0544322	1
Phytocomm	(Shi) Chang Pu	g006h0544023	1
Phytocomm	(Shi) Chang Pu	G006HS2700N1	1

- 848 spectra from 12 *Apo-Ident* customers from a total of 511 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	25 641
Type B	0	80	0	13 087
Type C	4	1	4	844

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.3953 % (> 98.8073 %)	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
Xiang Fu	8.99	—
Gui Zhi	12.21	—
Du Huo	12.42	—
Ze Xie	12.74	—
Chai Hu	13.07	—
Zi Wan	13.08	—
Qing Pi	13.23	—
She Gan	13.58	—
Gou Qi Zi	13.95	—
Sang Zhi	14.20	—
Lian Qiao	14.36	—
Lu Gen	14.46	—
Dang Gui Wei	15.16	—
Dang Gui	16.46	—
(Huai) Niu Xi	17.21	—
Gua Lou	17.31	—
Cang Zhu	18.29	—
Bai Mao Gen	18.51	—
Wei Ling Xian	18.51	—
Yi Mu Cao	18.75	—
Wu Yao	18.78	—
Chuan Xiong	20.87	—
Ku Shen	21.00	—
Huang Qi	21.57	—
Gan Jiang	21.87	—
San Leng	22.39	—
Di Gu Pi	22.51	—
Hou Po	22.79	—
Niu Bang Zi	22.91	—
Fang Feng	23.12	—
Tian Hua Fen	23.51	—
Ju Hua	23.82	—
Mu Dan Pi	24.15	—
Ji Li	24.28	—
Zhi Mu	24.60	—
Jiang Huang	25.29	—
Cang Er Zi	25.93	—
Chi Shao (Yao)	27.27	—
Huo Ma Ren	27.40	—
Chen Pi	27.75	—
Xie Bai	28.41	—
Sha Ren	28.89	—
Xian He Cao	28.97	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	29.15	—
Yuan Zhi	29.24	—
Jie Geng	29.38	—
Hong Jing Tian	30.23	—
Xu Duan	30.31	—
Shan Yao	30.39	—
Mang Xiao	30.87	—
(Bai) Dou Kou	30.89	—
Zhi Zi	31.02	—
Ba Ji Tian	31.41	—
Fu Zi	31.82	—
Yan Hu Suo	32.19	—
Dang Shen	32.58	—
E Zhu	32.59	—
Bai Shao Yao	32.94	—
Tu Fu Ling	33.12	—
Wu Wei Zi	33.40	—
Zi Su Zi	33.58	—
Pu Gong Ying	33.63	—
Chuan Lian Zi	33.87	—
Huang Bo	34.46	—
Long Dan (Cao)	34.57	—
Gan Cao	34.77	—
Huang Lian	34.84	—
He Huan Pi	35.40	—
Xin Yi	35.61	—
Mu Gua	35.84	—
Qiang Huo	35.89	—
Ma Bian Cao	36.80	—
Jin Yin Hua	37.44	—
Wu Zhu Yu	37.59	—
Sha Shen (Bei)	37.67	—
Yu Jin	37.68	—
Sheng Jiang	39.43	—
Dan Shen	39.48	—
Gu Sui Bu	39.68	—
Lai Fu Zi	39.71	—
Nü Zhen Zi	40.14	—
Bai Zhu	40.29	—
Qin Jiao	41.20	—
Shan Yu Rou	41.58	—
Ge Gen	42.39	—
He Shou Wu	43.31	—
Dong Gua Zi	43.57	—
Zi Hua Di Ding	45.02	—
Sang Ye	45.10	—
Hu Zhang	45.25	—
Mao Dong Qing	45.44	—
Huang Jing	45.60	—
Yin Yang Huo	45.66	—
Sang Bai Pi	46.18	—
Huang Bai	46.52	—
Sang Ji Shend	46.85	—
Du Zhong	47.06	—
Lian Zi	47.23	—
Fang Ji	47.50	—
Yin Chen Hao	47.97	—
Dan Dou Chi	48.03	—
Xi Xian Cao	48.37	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ji Xue Teng	48.82	—
Bai He	48.90	—
Bai Xian Pi	48.92	—
Gou Teng	49.69	—

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
62108	62108	0.00	9.35
62400	62400	0.00	8.99
62107	62107	0.00	9.28
62399	62399	0.00	9.91

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 2.1-2020-01

Validated substance/substance group **Ai Ye**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002258-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ai Ye; Artemisiae argyi folium

Special notes

When selecting the *Ai Ye* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ai Ye	1	0	3

Second-stage model

For differentiation of the substance/substance group *Ai 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 *Ai Ye*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ai Ye	G033H0613821	62771	40	from supplier
PhytoComm	Ai Ye	G033H0613821	62772	40	from supplier

Validation samples

A total of 39 821 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 *Ai Ye*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Ai 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	Ai Ye	G033H0613821	62771 [†]	20
PhytoComm	Ai Ye	G033H0613821	62772 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Ai Ye*.
- 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	Ai Ye	G033H0613022	1
Phytocomm	Ai Ye	G033H0613421	1
PhytoComm	Ai Ye	G033H0613021	1

- 850 spectra from 12 *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 *Ai 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 *Ai 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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

The substance/substance group *Ai Ye* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *Ai 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
Jue Ming Zi	17.41	—
Guang Huo Xiang	19.96	—
Shu Di (Huang)	22.48	—
Yu Xing Cao	23.98	—
Pu Gong Ying	29.40	—
(Sheng) Di Huang	36.58	—
Ye Jiao Teng	42.04	—
Wu Mei	43.39	—
Hua Shi	43.95	—
Jing Jie	49.31	—
Xuan Shen	49.76	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ai 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
62772	62772	0.00	17.74
62771	62771	0.00	17.41

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 2.1-2020-01

Validated substance/substance group **Ba Ji Tian**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60138-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G165H0445821	62407	40	from supplier
PhytoComm	Ba Ji Tian	G165H0445821	62408	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G165H0445821	62407 [†]	20
PhytoComm	Ba Ji Tian	G165H0445821	62408 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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
Dang Gui Wei	11.70	—
Ji Li	12.90	—
Chuan Xiong	12.91	—
Gua Lou	13.09	—
Chai Hu	14.04	—
Sang Zhi	14.07	—
Long Dan (Cao)	14.41	—
(Huai) Niu Xi	14.98	—
Dang Gui	15.47	—
Di Gu Pi	16.68	—
Jie Geng	16.80	—
Bai Mao Gen	17.50	—
San Leng	18.16	—
Chuan Lian Zi	18.63	—
She Gan	19.04	—
Fang Feng	19.39	—
Chen Pi	19.67	—
Gan Jiang	20.37	—
Zhi Mu	20.43	—
Gui Zhi	20.51	—
Gou Qi Zi	20.59	—
Lu Gen	20.69	—
Xie Bai	20.73	—
Cang Zhu	20.98	—
E Zhu	20.99	—
Du Huo	21.38	—
Sheng Jiang	23.34	—
Hou Po	23.66	—
Lai Fu Zi	23.70	—
Bai He	24.19	—
Mang Xiao	24.86	—
San Qi	24.97	—
Zi Wan	25.34	—
Bai Zhu	25.38	—
Mu Dan Pi	25.70	—
(Shi) Chang Pu	26.48	—
Huang Qi	26.48	—
Yi Mu Cao	27.54	—
Tian Hua Fen	27.99	—
Mu Gua	28.23	—
Huo Ma Ren	28.56	—
He Huan Pi	28.69	—
Zi Su Zi	28.81	—
Lian Qiao	28.88	—
Jiang Huang	29.02	—
Tu Fu Ling	29.11	—
Ju Hua	29.52	—
Sha Shen (Bei)	30.67	—
Qing Pi	30.93	—
Yan Hu Suo	31.20	—
Hong Jing Tian	31.37	—
Mai Men Dong	31.42	—
Xiang Fu	31.62	—
Xian He Cao	32.07	—
Bai Bian Dou	32.44	—
Wu Yao	32.60	—
Yuan Zhi	32.61	—
Dang Shen	32.92	—
Mai Ya	33.24	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Shao Yao	34.96	—
Shan Yao	35.13	—
Qin Jiao	36.45	—
Niu Bang Zi	36.56	—
Sha Ren	38.04	—
Da Zao	39.15	—
(Bai) Dou Kou	39.66	—
Cang Er Zi	40.32	—
Zhu Ling	41.63	—
Lian Zi	41.73	—
Rou Gui	42.05	—
Huang Jing	43.42	—
Mo Yao	43.73	—
Jin Yin Hua	44.34	—
Ku Shen	44.39	—
Wang Bu Liu Xing	44.67	—
Yu Jin	44.91	—
Zhi Ke	44.99	—
Xu Duan	45.21	—
Dong Gua Zi	45.50	—
Chi Shao (Yao)	45.85	—
Fu Ling	46.22	—
Huang Lian	46.48	—
Qiang Huo	46.54	—
Bai Xian Pi	47.00	—
Wu Wei Zi	47.29	—
Xin Yi	47.44	—
Ma Bian Cao	47.61	—
Gu Sui Bu	48.72	—
Sang Ji Shend	49.04	—
Ban Xia (Jiang)	49.43	—
Ji Xue Teng	49.61	—
Ban Zhi Lian	50.00	—

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
62408	62408	0.00	12.31
62407	62407	0.00	11.70

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 2.1-2020-01

Validated substance/substance group **Bai Bian Dou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60204-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Bai Bian Dou; Dolichoris lablab semen

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Bai Bian Dou	2	0	1

Second-stage model

For differentiation of the substance/substance group *Bai Bian Dou* 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 Bian Dou*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Bian Dou	G093HS177QM1	62209	40	from supplier
PhytoComm	Bai Bian Dou	G093HS177QM1	62210	40	from supplier
PhytoComm	Bai Bian Dou	G093H0510821	62527	40	from supplier
PhytoComm	Bai Bian Dou	G093H0510821	62528	40	from supplier

Validation samples

A total of 39 821 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 Bian Dou*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 Bian Dou*.
- 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 Bian Dou	G093HS177QM1	62209 [†]	20
PhytoComm	Bai Bian Dou	G093HS177QM1	62210 [†]	20
PhytoComm	Bai Bian Dou	G093H0510821	62527 [†]	20
PhytoComm	Bai Bian Dou	G093H0510821	62528 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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 Bian Dou*.
- 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 Bian Dou	g093h0510221	1
Phytocomm	Bai Bian Dou	G093H0510221	1

- 851 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 *Bai Bian Dou* 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 Bian Dou* 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	25 640
Type B	7	80	0	13 080
Type C	0	0	2	851

The substance/substance group *Bai Bian Dou* 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	99.9322 % (> 99.9033 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 Bian Dou* 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	5.23	—
Lian Qiao	6.11	—
Lian Zi	6.51	—
Shen Qu	6.92	—
Bai Zi Ren	7.50	—
Shan Yao	7.60	—
Ban Xia (Jiang)	7.82	—
Huo Ma Ren	9.26	—
Yuan Zhi	9.31	—
Ling Zhi	9.68	—
Fu Zi	11.00	—
Pi Pa Ye	11.72	—
Ren Dong Teng	11.73	—
He Huan Pi	11.92	—
Zhe Bei Mu	12.93	—
Fu Xiao Mai	12.96	—
Tai Zi Shen	13.06	—
Bai Shao Yao	13.25	—
Tian Hua Fen	13.33	—
Sheng Jiang	13.51	—
Rou Gui	14.20	—
Cang Er Zi	14.72	—
Gua Lou	15.56	—
Di Gu Pi	15.78	—
Zi Su Zi	16.14	—
Gui Zhi	16.46	—
Fu Ling	16.48	—
Hou Po	16.65	—
Chuan Xiong	16.78	—
Mao Dong Qing	17.39	—
Tao Ren	17.56	—
Jiao Gu Lan	17.68	—
San Qi	17.87	—
Yu Jin	17.93	—
Bai Xian Pi	17.97	—
Tu Fu Ling	18.95	—
Dan Shen	18.95	—
Chai Hu	19.03	—
Cang Zhu	19.17	—
Ji Xue Teng	19.47	—
Zhu Ru	19.50	—
Suan Zao Ren	19.52	—
Gou Teng	19.85	—
Yan Hu Suo	19.90	—
Jin Yin Hua	20.15	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chen Pi	20.35	—
Wu Wei Zi	20.93	—
Fang Ji	20.97	—
Ci Wu Jia	21.32	—
Che Qian Zi	21.46	—
Ye Jiao Teng	21.49	—
Xie Bai	21.58	—
Dan Dou Chi	21.82	—
Yi Yi Ren	22.17	—
Mo Yao	23.32	—
Yin Yang Huo	23.70	—
Yu Zhu	24.74	—
Dang Shen	24.89	—
Sang Zhi	25.79	—
She Gan	26.03	—
Dang Gui	26.33	—
Mang Xiao	26.39	—
Shan Yu Rou	26.66	—
Mu Dan Pi	26.68	—
Gou Qi Zi	26.86	—
Wang Bu Liu Xing	27.03	—
Dang Gui Wei	27.67	—
Zhi Ke	27.84	—
Zhi Mu	27.91	—
Ma Huang Gen	27.94	—
Lai Fu Zi	28.88	—
Qing Pi	28.97	—
E Zhu	29.10	—
Hong Jing Tian	29.55	—
Ma Huang	29.61	—
Gan Jiang	30.22	—
Dong Gua Zi	30.85	—
Qiang Huo	31.35	—
Du Zhong	31.62	—
Jiang Huang	31.75	—
Da Zao	31.90	—
Niu Bang Zi	32.06	—
Tian Men Dong	34.18	—
Chuang Mu Xiang	34.61	—
Huang Qin	34.70	—
Fo Shou	35.06	—
Ba Ji Tian	35.68	—
Bai He	36.07	—
Long Dan (Cao)	36.10	—
Long Yan Rou	36.15	—
Jie Geng	36.67	—
Huang Lian	37.15	—
Chuan Lian Zi	38.49	—
Fu Shen	39.30	—
Zi Hua Di Ding	39.64	—
Xiang Fu	39.70	—
Sha Ren	39.90	—
Fang Feng	40.50	—
Mu Gua	41.51	—
Ze Xie	42.15	—
(Huai) Niu Xi	42.20	—
Bai Zhu	42.57	—
Sha Shen (Bei)	42.64	—
Mai Men Dong	42.90	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Du Huo	43.04	—
Huang Qi	44.45	—
Gan Cao	44.97	—
(Bai) Dou Kou	45.18	—
Chi Shao (Yao)	45.84	—
(Shi) Chang Pu	46.17	—
Ku Shen	46.75	—
Wei Ling Xian	48.66	—
Yi Mu Cao	49.92	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Bian Dou* 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
62210	62210	0.00	6.51
62528	62528	0.00	14.72
62527	62527	0.00	15.29
62209	62209	0.00	5.23

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 2.1-2020-01

Validated substance/substance group **Bai He**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60093-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Bai He; Lili bulb

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G139H0607622	62137	40	from supplier
PhytoComm	Bai He	G139H0607622	62138	40	from supplier
PhytoComm	Bai He	G139H0607921	62739	40	from supplier
PhytoComm	Bai He	G139H0607921	62740	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G139H0607622	62137 [†]	20
PhytoComm	Bai He	G139H0607622	62138 [†]	20
PhytoComm	Bai He	G139H0607921	62739 [†]	20
PhytoComm	Bai He	G139H0607921	62740 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G139H0607421	2
PhytoComm	Bai He	G139H0607221	1
PhytoComm	Bai He	G139H0607421	1

- 848 spectra from 12 *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 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	25 641
Type B	0	80	0	13 087
Type C	0	0	5	848

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8240 %)	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
Chuan Lian Zi	9.01	—
Mai Ya	14.00	—
Mai Men Dong	14.51	—
Jie Geng	14.79	—
Di Gu Pi	14.89	—
Da Zao	16.75	—
San Qi	17.27	—
Chuan Xiong	17.65	—
Sang Zhi	19.85	—
Sheng Jiang	20.29	—
Long Dan (Cao)	22.49	—
Dang Gui	22.51	—
Lai Fu Zi	22.67	—
Mang Xiao	22.86	—
Bai Zhu	24.00	—
Qin Jiao	26.54	—
Tian Hua Fen	26.62	—
Mu Dan Pi	26.72	—
Mo Yao	26.76	—
Chai Hu	26.94	—
Dang Shen	27.38	—
Bai Bian Dou	27.64	—
E Zhu	28.21	—
San Leng	28.37	—
She Gan	29.35	—
Du Huo	30.17	—
Mu Gua	31.31	—
Ba Ji Tian	31.55	—
Zhi Mu	31.98	—
Ban Zhi Lian	32.69	—
Dang Gui Wei	33.23	—
Gan Jiang	34.41	—
Zhu Ling	34.51	—
Huang Qi	34.62	—
Gui Zhi	35.21	—
Yan Hu Suo	35.36	—
Lian Zi	35.90	—
Ban Xia (Jiang)	36.03	—
Rou Gui	37.12	—
Cang Zhu	37.27	—
Jiang Huang	38.01	—
Ji Li	38.73	—
Cang Er Zi	39.63	—
Zi Su Zi	40.28	—
Yi Yi Ren	40.48	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Qi Zi	40.86	—
Zhu Ru	41.18	—
(Huai) Niu Xi	41.29	—
Fu Ling	41.54	—
Qing Pi	42.65	—
Shan Yao	42.75	—
Tu Fu Ling	43.61	—
Xie Bai	43.70	—
Chen Pi	43.93	—
Fang Feng	44.98	—
Huo Ma Ren	45.26	—
Ma Huang Gen	45.71	—
Lu Gen	46.06	—
Tian Men Dong	46.84	—
(Shi) Chang Pu	47.26	—
He Huan Pi	47.76	—
Gua Lou	48.54	—
Sha Shen (Bei)	48.76	—
Bai Shao Yao	48.78	—
Sha Ren	48.99	—
Bai Xian Pi	49.80	—

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	9.37
62137	62137	0.00	9.27
62138	62138	0.00	9.01
62740	62740	0.00	9.38

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 2.1-2020-01

Validated substance/substance group **Bai Hua She She Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60039-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Bai Hua She She Cao	1	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

Validation samples

A total of 39 821 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 Hua She She Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 846 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	7	846

The substance/substance group *Bai Hua She She 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8232 %)	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 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
Han Lian Cao	9.59	—
Xia Ku Cao	9.78	—
Zhi Shi	15.81	—
Hu Zhang	15.88	—
Sang Ye	16.64	—
Xi Xian Cao	17.24	—
Gui Zhi	20.07	—
Pu Gong Ying	21.38	—
Jiang Huang	22.12	—
Huang Lian	22.99	—
Tu Fu Ling	24.57	—
Xin Yi	25.02	—
Yin Yang Huo	25.93	—
Sha Ren	26.23	—
Shan Yao	27.02	—
He Huan Pi	28.55	—
Rou Cong Rong	29.71	—
Nü Zhen Zi	29.99	—
Yi Mu Cao	30.79	—
Fu Zi	31.02	—
Yin Chen Hao	33.52	—
Cang Er Zi	34.04	—
Gu Sui Bu	34.68	—
Bu Gu Zhi	35.24	—
Xuan Shen	35.35	—
Huang Bo	35.71	—
Qiang Huo	35.82	—
Dan Zhu Ye	38.33	—
Shen Qu	38.71	—
Gou Teng	38.72	—
Dan Shen	42.01	—
Wei Ling Xian	42.46	—
Sang Ji Shend	42.88	—
Qing Pi	44.30	—
Du Zhong	45.68	—
Wu Zhu Yu	45.70	—
Mang Xiao	45.97	—
Chen Pi	46.96	—
Sang Bai Pi	47.42	—
Wu Yao	47.71	—
Jing Jie	48.88	—
Ju Hua	49.02	—
Xu Duan	50.02	—

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	9.59
62576	62576	0.00	9.98

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 2.1-2020-01

Validated substance/substance group **Bai Mao Gen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60654-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Bai Mao Gen; Imperatae cylindricae rhizoma

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Bai Mao Gen	1	0	1

Second-stage model

For differentiation of the substance/substance group *Bai Mao 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 *Bai Mao Gen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Bai Mao Gen	G125H0512821	62537	40	from supplier
PhytoComm	Bai Mao Gen	G125H0512821	62538	40	from supplier

Validation samples

A total of 39 821 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 Mao Gen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Mao 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	Bai Mao Gen	G125H0512821	62537 [†]	20
PhytoComm	Bai Mao Gen	G125H0512821	62538 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Bai Mao 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	Bai Mao Gen	G125H0512221	2

- 851 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 *Bai Mao 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 *Bai Mao 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	25 721
Type B	0	40	0	13 127
Type C	1	0	2	850

The substance/substance group *Bai Mao 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	100.0000 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.8837 % (> 99.2976 %)	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 Mao 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
Dang Gui Wei	7.11	—
Chai Hu	7.73	—
Gou Qi Zi	8.76	—
She Gan	10.70	—
Jie Geng	10.92	—
Huang Qi	14.90	—
Cang Zhu	15.07	—
Sang Zhi	15.47	—
Dang Gui	15.99	—
San Leng	16.08	—
(Huai) Niu Xi	17.03	—
Bai Zhu	17.05	—
Chuan Xiong	17.09	—
Zi Wan	17.25	—
Lu Gen	17.86	—
Gua Lou	18.02	—
Zhi Mu	19.71	—
Du Huo	19.84	—
Bai Shao Yao	20.87	—
Fang Feng	20.98	—
Long Dan (Cao)	21.38	—
Yuan Zhi	21.42	—
Ba Ji Tian	21.98	—
Mu Gua	22.51	—
Mu Dan Pi	22.78	—
Dang Shen	23.19	—
Gui Zhi	23.88	—
Di Gu Pi	25.14	—
Gan Jiang	25.56	—
Ji Li	26.93	—
Ju Hua	27.01	—
Qing Pi	28.08	—
Hong Jing Tian	28.34	—
(Shi) Chang Pu	29.04	—
Chen Pi	29.76	—
Xiang Fu	29.91	—
Lian Qiao	30.09	—
Mang Xiao	30.11	—
Bai He	30.14	—
Tian Hua Fen	30.58	—
Xie Bai	31.28	—
Chi Shao (Yao)	31.93	—
E Zhu	32.07	—
Chuan Lian Zi	32.81	—
Yi Mu Cao	33.19	—
Hou Po	33.26	—
Lai Fu Zi	33.51	—
Wu Yao	34.41	—
Qin Jiao	34.41	—
Mai Men Dong	34.65	—
Huo Ma Ren	35.02	—
Zhi Ke	36.53	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	37.36	—
Sheng Jiang	37.85	—
Ku Shen	38.13	—
Zi Su Zi	39.25	—
San Qi	39.88	—
Jiang Huang	41.09	—
Sha Shen (Bei)	41.22	—
Shan Yao	42.36	—
Xian He Cao	42.79	—
Ze Xie	42.88	—
He Huan Pi	42.99	—
Jin Yin Hua	43.44	—
Niu Bang Zi	44.25	—
Lian Zi	44.44	—
Da Zao	44.69	—
Wei Ling Xian	45.20	—
Gu Sui Bu	46.73	—
Cang Er Zi	47.13	—
Long Yan Rou	47.81	—
Bai Xian Pi	47.88	—
Mai Ya	47.97	—
Ban Zhi Lian	48.07	—
Huang Jing	48.36	—
Wu Wei Zi	48.41	—
Yan Hu Suo	48.61	—
Bai Bian Dou	48.91	—
Ji Xue Teng	49.13	—
Huang Lian	49.48	—
Sha Ren	49.49	—
Rou Gui	49.82	—
Dong Gua Zi	50.05	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Bai Mao 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
62537	62537	0.00	7.80
62538	62538	0.00	7.11

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 2.1-2020-01

Validated substance/substance group **Bai Shao Yao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60007-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G179H0502621	61955	40	from supplier
PhytoComm	Bai Shao Yao	G179H0502621	61956	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 521 spectra from a total of 315 batches from further 172 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	G179H0502621	61955 [†]	20
PhytoComm	Bai Shao Yao	G179H0502621	61956 [†]	20
PhytoComm	Bai Shao Yao	G179HS065RQ1	62275 [†]	20
PhytoComm	Bai Shao Yao	G179HS065RQ1	62276	59
PhytoComm	Bai Shao Yao	G179HS065RT1	62435 [†]	20

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Bai Shao Yao	G179HS065RT1	62436 [†]	20
PhytoComm	Bai Shao Yao	G179HS065RW1	62601 [†]	20
PhytoComm	Bai Shao Yao	G179HS065RW1	62602 [†]	20

- 12 968 spectra from a total of 316 batches from further 172 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	G179H0502322	1
Phytocomm	Bai Shao Yao	G179H0502121	1
Phytocomm	Bai Shao Yao	g179h0502322	1
Phytocomm	Bai Shao Yao	g17h0502323	1
Phytocomm	Bai Shao Yao	G179H0502521	5
PhytoComm	Bai Shao Yao	1214401	1
PhytoComm	Bai Shao Yao	G179H0502322	1

- 842 spectra from 12 *Apo-Ident* customers from a total of 510 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	0	279	1	25 521
Type B	1	199	0	12 967
Type C	1	3	8	841

[†]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	100.0000 % (> 99.9701 %)	99.6429 % (> 98.5714 %)
Type B	99.9952 % (> 99.9662 %)	100.0000 % (> 96.9849 %)
Type C	99.5349 % (> 98.9462 %)	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
She Gan	5.84	—
Tai Zi Shen	6.26	—
Rou Gui	7.76	—
Pi Pa Ye	7.87	—
Ye Jiao Teng	7.98	—
Di Gu Pi	8.17	—
He Huan Pi	8.42	—
Zhe Bei Mu	8.61	—
Lian Zi	8.95	—
Gua Lou	9.10	—
Fu Xiao Mai	9.31	—
Ban Xia (Jiang)	9.41	—
Bai Bian Dou	9.56	—
Ji Li	9.71	—
Shen Qu	9.79	—
Zhu Ru	9.94	—
Yuan Zhi	10.24	—
Shan Yao	11.07	—
Fo Shou	11.80	—
Ci Wu Jia	12.13	—
Tao Ren	12.23	—
Bai Zi Ren	12.31	—
Huang Qi	12.37	—
Gui Zhi	12.75	—
Sheng Jiang	12.81	—
Tian Hua Fen	13.01	—
Ji Xue Teng	13.46	—
Gou Qi Zi	13.73	—
Chuan Xiong	13.82	—
Lian Qiao	13.96	—
Dang Shen	14.11	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Yi Ren	14.74	—
Fu Ling	15.45	—
Chen Pi	16.44	—
Huo Ma Ren	16.48	—
Gou Teng	16.52	—
Da Zao	18.13	—
Dang Gui Wei	18.58	—
Ling Zhi	18.70	—
Shan Yu Rou	19.00	—
Du Huo	19.08	—
Mu Dan Pi	19.21	—
San Qi	19.26	—
San Leng	19.26	—
Ren Dong Teng	19.32	—
Wang Bu Liu Xing	19.38	—
Tian Men Dong	19.86	—
Mu Gua	20.04	—
Long Yan Rou	20.08	—
(Huai) Niu Xi	20.40	—
Jie Geng	21.04	—
Fu Zi	21.27	—
Zhi Mu	21.61	—
Bai Xian Pi	21.97	—
Jin Yin Hua	22.09	—
Dang Gui	22.20	—
Bai Mao Gen	22.21	—
Yu Zhu	22.27	—
Chi Shao (Yao)	22.55	—
Bai Zhu	22.83	—
Zhi Ke	23.89	—
Yan Hu Suo	23.90	—
Sang Zhi	24.16	—
Long Dan (Cao)	24.42	—
Jiao Gu Lan	24.79	—
Qin Jiao	24.80	—
Chai Hu	25.16	—
Tu Fu Ling	25.20	—
Ma Huang	25.45	—
Cang Zhu	25.69	—
Fu Shen	25.96	—
Huang Qin	26.27	—
Zi Wan	26.88	—
Wu Wei Zi	26.93	—
Hou Po	27.24	—
Ma Huang Gen	27.42	—
Fang Ji	27.70	—
Lu Gen	28.85	—
Yu Jin	29.41	—
Mao Dong Qing	29.61	—
Dan Shen	29.61	—
Che Qian Zi	30.09	—
Yin Yang Huo	30.57	—
Mang Xiao	33.21	—
Suan Zao Ren	33.28	—
Dan Dou Chi	34.00	—
Mai Men Dong	36.00	—
Qing Pi	36.07	—
(Shi) Chang Pu	36.36	—
Ze Xie	38.49	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Cao	38.72	—
Xiang Fu	39.01	—
Ba Ji Tian	39.16	—
Hong Jing Tian	39.91	—
Gan Jiang	40.31	—
Fang Feng	40.36	—
Wu Yao	40.88	—
Sha Shen (Bei)	41.00	—
Ban Zhi Lian	41.98	—
Chuang Mu Xiang	44.12	—
Ju Hua	44.53	—
Bai He	44.82	—
Qiang Huo	45.18	—
Huang Lian	45.19	—
Lai Fu Zi	46.16	—
E Zhu	48.59	—
Xian He Cao	48.85	—
Du Zhong	49.33	—
Chuan Lian Zi	49.62	—

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
62435	62435	0.00	8.58
61956	61956	0.00	5.84
62275	62275	0.00	6.26
62436	62436	0.00	7.87
62602	62602	0.00	7.76
62601	62601	0.00	8.30
61955	61955	0.00	5.94

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 2.1-2020-01

Validated substance/substance group **Bai Xian Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60208-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	4	156	4	25 637
Type B	6	79	1	13 081
Type C	1	0	1	851

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	99.9709 % (> 99.9560 %)	97.5000 % (> 95.6250 %)
Type B	99.9419 % (> 99.9130 %)	98.7500 % (> 95.0000 %)
Type C	99.9225 % (> 99.3397 %)	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	3.54	—
Ye Jiao Teng	5.48	—
Tu Fu Ling	6.02	—
Ling Zhi	7.25	—
Mao Dong Qing	8.32	—
Yin Yang Huo	8.81	—
Ren Dong Teng	10.34	—
Yu Jin	10.46	—
Bai Bian Dou	10.51	—
Lian Zi	11.00	—
Hou Po	11.12	—
Yuan Zhi	12.08	—
Shen Qu	13.42	—
Qiang Huo	13.90	—
Rou Gui	14.16	—
Chai Hu	14.98	—
Ji Xue Teng	15.41	—
Che Qian Zi	16.19	—
He Huan Pi	17.68	—
Gou Teng	17.77	—
Dan Shen	18.02	—
Yan Hu Suo	18.91	—
Ban Xia (Jiang)	19.27	—
Ma Huang Gen	19.30	—
Ji Li	20.12	—
Du Zhong	20.34	—
Lian Qiao	20.62	—
Shan Yao	20.64	—
Jiao Gu Lan	21.07	—
Fang Ji	21.07	—
Pi Pa Ye	21.13	—
Huo Ma Ren	21.27	—
Sheng Jiang	21.43	—
Jin Yin Hua	21.51	—
Ma Huang	21.82	—
Chen Pi	22.23	—
Chuan Xiong	22.33	—
Gui Zhi	22.38	—
Qing Pi	23.16	—
Bai Shao Yao	23.82	—
Fu Xiao Mai	24.24	—
Mu Dan Pi	24.39	—
Zhi Ke	24.58	—
Yi Yi Ren	24.90	—
Zi Hua Di Ding	25.59	—
Hong Jing Tian	25.69	—
Tao Ren	25.90	—
Di Gu Pi	26.44	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Zi	26.66	—
Gua Lou	26.66	—
Huang Lian	27.62	—
Tian Hua Fen	27.92	—
Zhe Bei Mu	28.84	—
Fu Ling	28.98	—
Tai Zi Shen	29.57	—
Wu Wei Zi	29.81	—
Long Yan Rou	30.71	—
Cang Zhu	31.66	—
Ci Wu Jia	31.78	—
Fu Shen	32.16	—
Ban Zhi Lian	32.65	—
Gou Qi Zi	32.85	—
Fo Shou	33.02	—
Da Zao	34.47	—
Mang Xiao	34.96	—
Huang Bai	35.39	—
Tian Men Dong	35.45	—
Bai Zi Ren	37.48	—
Sang Zhi	37.68	—
Zhu Ru	38.68	—
Yu Zhu	39.48	—
Suan Zao Ren	39.86	—
Pu Gong Ying	40.10	—
San Qi	40.59	—
Jie Geng	40.87	—
Chi Shao (Yao)	41.67	—
Wang Bu Liu Xing	42.10	—
Dang Gui Wei	42.64	—
Ze Xie	42.87	—
Zhi Mu	43.45	—
Huang Qin	43.66	—
Shan Yu Rou	43.74	—
Bai He	43.79	—
Dang Shen	44.15	—
Chuang Mu Xiang	44.96	—
Dang Gui	45.96	—
Bai Zhu	46.10	—
Chuan Lian Zi	46.52	—
Zi Su Zi	47.41	—
Cang Er Zi	48.15	—
Jiang Huang	48.94	—
Xi Xian Cao	49.44	—

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
62450	62450	0.00	4.54
62449	62449	0.00	4.71
62286	62286	0.00	3.64
62285	62285	0.00	3.54

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 2.1-2020-01

Validated substance/substance group **Bai Zhu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60015-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Bai Zhu	4	0	2

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	G041H0501621	61837	40	from supplier
PhytoComm	Bai Zhu	G041H0501721	62219	40	from supplier
PhytoComm	Bai Zhu	G041H0501721	62220	40	from supplier
PhytoComm	Bai Zhu	G041H0501822	62385	40	from supplier
PhytoComm	Bai Zhu	G041H0501822	62386	40	from supplier
PhytoComm	Bai Zhu	G041HS064SG1	62651	40	from supplier
PhytoComm	Bai Zhu	G041HS064SG1	62652	40	from supplier

Validation samples

A total of 39 821 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 Zhu*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 25 521 spectra from a total of 315 batches from further 172 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 *Type B* samples 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.

- 140 spectra of 7 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	G041H0501621	61837 [†]	20
PhytoComm	Bai Zhu	G041H0501721	62219 [†]	20
PhytoComm	Bai Zhu	G041H0501721	62220 [†]	20
PhytoComm	Bai Zhu	G041H0501822	62385 [†]	20
PhytoComm	Bai Zhu	G041H0501822	62386 [†]	20
PhytoComm	Bai Zhu	G041HS064SG1	62651 [†]	20
PhytoComm	Bai Zhu	G041HS064SG1	62652 [†]	20

- 13 027 spectra from a total of 316 batches from further 172 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 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 Zhu	G041H0501221	1
Phytocomm	Bai Zhu	g041h0501221	3
Phytocomm	Bai Zhu	G041H0501422	3
PhytoComm	Bai Zhu	G041H0501621	1

- 845 spectra from 12 *Apo-Ident* customers from a total of 512 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	4	277	3	25 517
Type B	8	136	4	13 019
Type C	1	1	7	844

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	99.9903 % (> 99.9754 %)	98.9286 % (> 97.8571 %)
Type B	99.9612 % (> 99.9323 %)	97.1429 % (> 95.0000 %)
Type C	99.9483 % (> 99.3598 %)	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 *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	3.41	—
Jiao Gu Lan	8.87	—
Jin Yin Hua	10.30	—
Mu Dan Pi	11.05	—
Hong Jing Tian	11.18	—
Dang Shen	12.33	—
Huang Qi	12.86	—
Di Gu Pi	13.00	—
Shan Yu Rou	13.11	—
Shan Yao	13.61	—
Chuan Xiong	14.13	—
Gou Qi Zi	14.19	—
Sang Zhi	15.04	—
Dang Gui Wei	15.41	—
San Leng	15.52	—
Dan Shen	15.69	—
Mu Gua	16.19	—
Yuan Zhi	16.80	—
Chen Pi	17.06	—
Wu Wei Zi	17.22	—
Zi Hua Di Ding	17.64	—
Mai Men Dong	17.84	—
Zhi Ke	17.94	—
Cang Zhu	17.95	—
Gua Lou	18.43	—
Pi Pa Ye	18.58	—
Chuang Mu Xiang	19.00	—
She Gan	19.17	—
San Qi	19.18	—
Suan Zao Ren	19.56	—
Bai He	19.76	—
Qing Pi	19.89	—
Lian Qiao	19.94	—
Chuan Lian Zi	20.07	—
Bai Bian Dou	20.36	—
Chai Hu	20.97	—
Fang Ji	21.15	—
Zhi Mu	22.42	—
Hong Hua	22.57	—
Ji Li	22.59	—
Hou Po	22.68	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Dong Teng	22.72	—
Gan Cao	22.98	—
Yu Zhu	23.38	—
Mao Dong Qing	23.41	—
Long Dan (Cao)	23.59	—
Bai Shao Yao	23.70	—
Bai Mao Gen	23.90	—
Qin Jiao	24.47	—
Gui Zhi	24.54	—
Lian Zi	24.61	—
Zhe Bei Mu	25.15	—
Cang Er Zi	25.76	—
Chi Shao (Yao)	25.91	—
Yin Yang Huo	26.33	—
Yan Hu Suo	26.37	—
Long Yan Rou	26.89	—
Fu Zi	26.91	—
Du Huo	26.95	—
Ye Jiao Teng	26.98	—
Tian Hua Fen	27.00	—
Lu Gen	27.36	—
Bai Xian Pi	27.51	—
Huang Bai	27.64	—
(Huai) Niu Xi	27.78	—
Lai Fu Zi	27.81	—
Ma Huang	27.91	—
Sheng Jiang	28.26	—
Shen Qu	28.34	—
Qiang Huo	29.25	—
Mang Xiao	29.32	—
Tai Zi Shen	29.83	—
Dan Dou Chi	29.83	—
Rou Gui	29.88	—
Gan Jiang	29.97	—
Dang Gui	30.20	—
Huang Qin	30.24	—
He Huan Pi	30.55	—
Ji Xue Teng	30.63	—
Sang Ye	30.66	—
Du Zhong	30.91	—
Tu Fu Ling	31.04	—
Bai Zi Ren	31.62	—
Zi Wan	31.91	—
Ling Zhi	32.45	—
(Shi) Chang Pu	32.72	—
Gou Teng	33.65	—
Da Zao	33.89	—
Tao Ren	34.31	—
Xiang Fu	34.46	—
Tian Men Dong	34.61	—
Huang Lian	34.86	—
Ba Ji Tian	35.50	—
Yi Mu Cao	35.66	—
Ban Zhi Lian	36.32	—
Ze Xie	37.33	—
Fu Xiao Mai	37.65	—
Ci Wu Jia	37.81	—
Ban Xia (Jiang)	37.82	—
Fu Ling	37.90	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huo Ma Ren	38.28	—
E Zhu	38.43	—
Yu Jin	39.34	—
Pu Gong Ying	39.70	—
Ma Huang Gen	39.74	—
Che Qian Zi	41.03	—
Fang Feng	41.49	—
Fo Shou	41.91	—
Niu Bang Zi	41.96	—
Yi Yi Ren	42.31	—
Zi Su Zi	42.90	—
Zhu Ru	43.83	—
Sha Shen (Bei)	44.26	—
Xie Bai	44.74	—
Mai Ya	44.87	—
Ku Shen	46.43	—
Wu Mei	46.73	—
Xi Xian Cao	48.06	—
Wu Yao	48.21	—
Mo Yao	48.74	—
Jiang Huang	49.08	—
Fu Shen	49.79	—

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
61837	61837	0.00	6.27
62385	62385	0.00	4.96
62219	62219	0.00	3.41
62652	62652	0.00	8.87
62220	62220	0.00	4.06
62651	62651	0.00	9.22
62386	62386	0.00	5.28

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 2.1-2020-01

Validated substance/substance group **Bai Zi Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60197-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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	g952f110315	1
Phytocomm	Bai Zi Ren	g052f110315	1
Phytocomm	Bai Zi Ren	G052HS187	2

- 849 spectra from 12 *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 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	25 721
Type B	0	37	3	13 127
Type C	0	0	4	849

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8246 %)	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	6.05	—
Ban Xia (Jiang)	8.13	—
Bai Bian Dou	8.16	—
Ji Li	8.42	—
Ling Zhi	10.39	—
Bai Shao Yao	10.70	—
Rou Gui	10.91	—
Shen Qu	10.91	—
Fu Xiao Mai	11.06	—
He Huan Pi	11.55	—
Huo Ma Ren	11.61	—
Tai Zi Shen	12.61	—
Di Gu Pi	13.12	—
Sheng Jiang	13.26	—
Gui Zhi	13.90	—
Yuan Zhi	14.41	—
Ci Wu Jia	14.63	—
Tian Hua Fen	14.65	—
Shan Yao	14.84	—
Gua Lou	15.47	—
Zhe Bei Mu	15.60	—
Pi Pa Ye	15.67	—
Ren Dong Teng	16.51	—
Zhu Ru	16.59	—
Tao Ren	16.69	—
Tu Fu Ling	17.09	—
Bai Xian Pi	17.09	—
San Qi	17.23	—
Chen Pi	17.43	—
Chuan Xiong	17.55	—
Lian Qiao	17.66	—
Ye Jiao Teng	18.30	—
Dang Shen	18.95	—
Tian Men Dong	19.18	—
Fu Ling	19.30	—
Yi Yi Ren	19.48	—
Mao Dong Qing	20.00	—
Dang Gui Wei	20.69	—
Gou Teng	20.76	—
Dan Shen	22.27	—
Wang Bu Liu Xing	22.65	—
Fu Zi	22.93	—
Yu Zhu	22.97	—
Fo Shou	22.98	—
Suan Zao Ren	23.68	—
Yu Jin	23.90	—
Cang Zhu	24.11	—
Jiao Gu Lan	25.22	—
Ji Xue Teng	25.30	—
Dang Gui	25.89	—
Chai Hu	25.89	—
Da Zao	26.64	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Yan Rou	26.67	—
Zhi Mu	27.06	—
Hou Po	27.12	—
Dan Dou Chi	27.28	—
Shan Yu Rou	27.89	—
Ma Huang Gen	28.16	—
Yan Hu Suo	30.08	—
Zhi Ke	30.14	—
Che Qian Zi	30.27	—
Gou Qi Zi	31.09	—
Jin Yin Hua	31.41	—
Hong Jing Tian	31.58	—
Yin Yang Huo	31.86	—
Fang Ji	32.63	—
Ma Huang	32.75	—
Wu Wei Zi	33.22	—
Fu Shen	36.08	—
Huang Qin	37.48	—
Mang Xiao	39.53	—
Qing Pi	40.14	—
Mu Dan Pi	40.35	—
Chuang Mu Xiang	40.80	—
Qiang Huo	41.30	—
Ze Xie	41.75	—
Chuan Lian Zi	42.32	—
Bai Zhu	43.15	—
Jie Geng	44.41	—
Sang Zhi	44.56	—
Du Zhong	45.02	—
Huang Lian	46.55	—
Gan Cao	48.87	—

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
62660	62660	0.00	6.05
62659	62659	0.00	6.31

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 2.1-2020-01

Validated substance/substance group **Ban Xia (Jiang)**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60019-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ban Xia (Jiang)	2	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)	G191HS090QL1	62201	40	from supplier
PhytoComm	Ban Xia (Jiang)	G191HS090QL1	62202	40	from supplier
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62599	40	from supplier
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62600	40	from supplier

Validation samples

A total of 39 821 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 *Ban Xia (Jiang)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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)	G191HS090QL1	62201 [†]	20
PhytoComm	Ban Xia (Jiang)	G191HS090QL1	62202 [†]	20
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62599 [†]	20
PhytoComm	Ban Xia (Jiang)	G191HS090SH1	62600 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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)	G191H0526322	1
Phytocomm	Ban Xia (Jiang)	g191h0526121	1

- 851 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 *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	3	159	1	25 638
Type B	9	77	3	13 078
Type C	0	0	2	851

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	99.9903 % (> 99.9754 %)	99.3750 % (> 97.5000 %)
Type B	99.9491 % (> 99.9202 %)	96.2500 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Sheng Jiang	4.16	—
Di Gu Pi	4.34	—
Rou Gui	4.69	—
Zhu Ru	5.42	—
Gui Zhi	6.43	—
Yi Yi Ren	7.14	—
Lian Zi	8.01	—
Bai Shao Yao	8.33	—
He Huan Pi	8.35	—
Fu Xiao Mai	8.70	—
Bai Bian Dou	8.86	—
Zhe Bei Mu	9.89	—
Ci Wu Jia	10.63	—
Shen Qu	10.95	—
Ji Li	11.85	—
Ji Xue Teng	11.90	—
Tao Ren	12.37	—
Fu Ling	12.50	—
Tai Zi Shen	12.60	—
Bai Zi Ren	13.08	—
Gou Teng	13.23	—
Ling Zhi	16.10	—
Tian Hua Fen	16.11	—
Huo Ma Ren	17.30	—
Gua Lou	17.31	—
Fo Shou	17.39	—
Ye Jiao Teng	17.49	—
Pi Pa Ye	18.76	—
Wang Bu Liu Xing	18.79	—
Da Zao	19.24	—
Bai Xian Pi	19.43	—
Yuan Zhi	19.50	—
Ma Huang Gen	19.59	—
Fu Shen	19.62	—
Long Yan Rou	20.02	—
Shan Yao	20.28	—
Tian Men Dong	21.28	—
Lian Qiao	21.82	—
Yan Hu Suo	22.95	—
Chen Pi	23.45	—
Yu Jin	25.62	—
Ren Dong Teng	26.55	—
Chuan Xiong	26.62	—
Dang Shen	27.20	—
Fu Zi	28.49	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Mu	28.52	—
Tu Fu Ling	28.72	—
Shan Yu Rou	29.58	—
Huang Qin	30.97	—
Zhi Ke	31.00	—
Dan Shen	31.50	—
Wu Wei Zi	31.68	—
Fang Ji	31.70	—
Jin Yin Hua	31.83	—
Jiao Gu Lan	32.41	—
Dang Gui Wei	32.95	—
Dan Dou Chi	33.72	—
Yu Zhu	33.72	—
Yin Yang Huo	34.34	—
Ma Huang	34.98	—
San Qi	35.20	—
Mao Dong Qing	35.40	—
Dang Gui	35.88	—
Hou Po	37.11	—
Che Qian Zi	38.46	—
Cang Zhu	39.11	—
Mang Xiao	39.43	—
Chai Hu	41.25	—
Bai Zhu	44.85	—
Suan Zao Ren	45.01	—
Sang Zhi	45.56	—
Huang Lian	46.16	—
Hong Jing Tian	46.28	—
Jie Geng	46.71	—
Gou Qi Zi	48.05	—
Ze Xie	48.10	—
Gan Cao	48.10	—
Qiang Huo	49.78	—
Mai Men Dong	50.00	—

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
62202	62202	0.00	4.62
62600	62600	0.00	6.04
62201	62201	0.00	4.16
62599	62599	0.00	6.12

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 2.1-2020-01

Validated substance/substance group **Ban Zhi Lian**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60038-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 849 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 *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	25 721
Type B	0	39	1	13 127
Type C	0	0	4	849

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8246 %)	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
Sheng Jiang	16.18	—
Bai He	16.71	—
Tu Fu Ling	18.84	—
Gui Zhi	19.50	—
Mai Ya	21.99	—
Sang Zhi	23.43	—
San Qi	24.23	—
Mang Xiao	25.25	—
Rou Gui	25.90	—
Fu Ling	26.11	—
Bai Xian Pi	27.65	—
Chuan Xiong	28.09	—
E Zhu	28.48	—
Lian Zi	28.57	—
Mu Dan Pi	28.63	—
Di Gu Pi	29.36	—
Ji Xue Teng	30.48	—
San Leng	30.53	—
Yan Hu Suo	30.74	—
Tian Hua Fen	31.74	—
Bai Zhu	32.36	—
Long Dan (Cao)	32.54	—
Zhu Ling	32.75	—
Ji Li	33.00	—
Gou Teng	33.31	—
Shan Yao	33.47	—
Jiang Huang	33.60	—
Ban Xia (Jiang)	33.68	—
Chuan Lian Zi	34.22	—
Ye Jiao Teng	35.24	—
She Gan	35.44	—
Jie Geng	35.73	—
Chai Hu	36.89	—
Dan Dou Chi	36.93	—
Bai Bian Dou	36.95	—
Bai Shao Yao	37.29	—
Lu Gen	37.43	—
Yuan Zhi	37.53	—
Yi Yi Ren	37.66	—
Fo Shou	37.93	—
Ling Zhi	38.16	—
Dang Gui	38.19	—
(Bai) Dou Kou	38.57	—
Zhu Ru	38.73	—
He Huan Pi	39.25	—
Fang Ji	39.31	—
Zhi Ke	39.34	—
Shen Qu	39.60	—
Dang Shen	39.61	—
Tian Men Dong	39.93	—
Ma Huang Gen	40.16	—
Mao Dong Qing	40.24	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Dong Teng	40.66	—
Mai Men Dong	40.67	—
Sha Shen (Bei)	40.82	—
Qin Jiao	41.44	—
Du Huo	42.06	—
Yu Jin	42.09	—
Zhi Mu	42.90	—
Sha Ren	42.91	—
Wu Wei Zi	43.16	—
Gan Jiang	43.69	—
Da Zao	44.09	—
Qiang Huo	44.15	—
(Shi) Chang Pu	44.44	—
Huo Ma Ren	44.54	—
Wang Bu Liu Xing	44.95	—
Qing Pi	45.18	—
Huang Lian	45.44	—
Xie Bai	45.83	—
Chen Pi	46.05	—
Ba Ji Tian	46.51	—
Zi Su Zi	46.59	—
Gou Qi Zi	46.66	—
Jin Yin Hua	46.99	—
Du Zhong	46.99	—
Tai Zi Shen	47.10	—
Cang Er Zi	47.23	—
Cang Zhu	47.36	—
Xiang Fu	47.47	—
Fu Shen	47.58	—
Huang Qi	47.64	—
Yi Mu Cao	47.71	—
Dang Gui Wei	47.92	—
Yin Yang Huo	48.67	—
Dan Shen	48.90	—
Lai Fu Zi	49.29	—
Fu Xiao Mai	49.51	—
(Huai) Niu Xi	49.59	—
Lian Qiao	49.65	—

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
62754	62754	0.00	16.49
62753	62753	0.00	16.18

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 2.1-2020-01

Validated substance/substance group **Bu Gu Zhi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60023-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G205H1247621	61939	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247621	61940	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247822	62487	40	from supplier
PhytoComm	Bu Gu Zhi	G205H1247822	62488	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G205H1247621	61939 [†]	20
PhytoComm	Bu Gu Zhi	G205H1247621	61940 [†]	20
PhytoComm	Bu Gu Zhi	G205H1247822	62487 [†]	20
PhytoComm	Bu Gu Zhi	G205H1247822	62488 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	2
Phytocomm	Bu Gu Zhi	g205h1247222	1
Phytocomm	Bu Gu Zhi	G205H1247421	2

- 848 spectra from 12 *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 *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	25 641
Type B	2	80	0	13 085
Type C	1	2	3	847

The substance/substance group *Bu Gu 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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	99.9709 % (> 99.9420 %)	100.0000 % (> 92.5000 %)
Type C	99.9419 % (> 99.3538 %)	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
Ma Bian Cao	6.77	—
Dan Zhu Ye	8.46	—
Yin Chen Hao	11.46	—
Sang Bai Pi	14.27	—
He Shou Wu	14.54	—
Jing Jie	17.32	—
Sha Ren	18.15	—
Du Zhong	18.24	—
Gu Sui Bu	18.45	—
Xu Duan	20.43	—
Gui Zhi	21.74	—
Xian He Cao	21.94	—
Sang Ji Shend	22.85	—
He Huan Pi	24.16	—
Tu Si Zi	24.85	—
Yu Xing Cao	25.22	—
Yin Yang Huo	25.26	—
Wu Zhu Yu	25.89	—
Yi Mu Cao	26.10	—
Nü Zhen Zi	28.05	—
(Sheng) Di Huang	28.98	—
Jin Qian Cao	30.14	—
Dan Shen	32.69	—
Mang Xiao	32.69	—
Wu Jia Pi	33.26	—
Fu Zi	33.97	—
Yu Jin	34.72	—
Huang Bo	35.36	—
Gou Teng	35.90	—
Zhi Shi	36.36	—
Qiang Huo	36.48	—
Ji Li	39.01	—
Xuan Shen	40.23	—
Ge Gen	40.69	—
Sang Ye	40.77	—
Xia Ku Cao	41.13	—
Hou Po	41.29	—
Chen Pi	41.40	—
Hu Zhang	41.52	—
Bai Hua She She Cao	41.95	—
Fang Feng	42.17	—
Pu Gong Ying	42.37	—
Jiang Huang	42.84	—
(Bai) Dou Kou	44.88	—
Wu Yao	45.08	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xin Yi	45.50	—
Han Lian Cao	46.13	—
Huang Jing	46.40	—
Hong Jing Tian	46.74	—
Hua Shi	47.48	—
Huang Lian	48.25	—

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
61939	61939	0.00	7.02
62487	62487	0.00	19.97
61940	61940	0.00	6.77
62488	62488	0.00	19.88

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 2.1-2020-01

Validated substance/substance group **Cang Er Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60054-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Cang Er Zi	3	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	G251H1401622	61897	40	from supplier
PhytoComm	Cang Er Zi	G251H1401622	61898	40	from supplier
PhytoComm	Cang Er Zi	G251H1401622	62155	40	from supplier
PhytoComm	Cang Er Zi	G251H1401622	62156	40	from supplier
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 39 821 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 *Cang Er Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 481 spectra from a total of 316 batches from further 172 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 *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	G251H1401622	61897 [†]	20
PhytoComm	Cang Er Zi	G251H1401622	61898 [†]	20
PhytoComm	Cang Er Zi	G251H1401622	62155 [†]	20
PhytoComm	Cang Er Zi	G251H1401622	62156 [†]	20

continued on the next page

continued from previous page

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

- 13 007 spectra from a total of 317 batches from further 172 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	G251H1401521	1
PhytoComm	Cang Er Zi	G251H1401322	1

- 849 spectra from 12 *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 *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	320	0	25 481
Type B	0	160	0	13 007
Type C	1	1	3	848

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,

[†]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 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 % (> 98.1250 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 96.2500 %)
Type C	99.5349 % (> 98.9472 %)	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 *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	9.97	—
Shan Yao	12.02	—
Bai Bian Dou	14.43	—
E Zhu	16.58	—
Sang Zhi	16.71	—
Ku Shen	17.71	—
Ji Li	18.40	—
Xie Bai	19.90	—
Zhi Zi	19.92	—
Wu Wei Zi	20.31	—
San Qi	20.55	—
Gan Jiang	20.55	—
Huo Ma Ren	20.68	—
Dong Gua Zi	21.15	—
Yi Mu Cao	21.49	—
Shen Qu	21.53	—
She Gan	22.80	—
Chai Hu	22.89	—
Gou Qi Zi	23.09	—
Di Gu Pi	23.09	—
Jiang Huang	23.15	—
Chuan Lian Zi	23.22	—
Wu Mei	23.24	—
Niu Bang Zi	23.41	—
Zhi Mu	24.47	—
Jin Yin Hua	24.69	—
Gui Zhi	24.73	—
Mang Xiao	24.78	—
Sha Shen (Bei)	24.81	—
Xiang Fu	25.24	—
Yan Hu Suo	25.44	—
Sheng Jiang	25.53	—
Ba Ji Tian	26.11	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Dan (Cao)	26.77	—
Qing Pi	27.06	—
Mo Yao	27.20	—
Lai Fu Zi	27.30	—
Wei Ling Xian	27.76	—
Gua Lou	27.86	—
Chuan Xiong	28.43	—
Sha Ren	29.27	—
Wang Bu Liu Xing	29.32	—
Qiang Huo	29.94	—
Fang Ji	30.01	—
Chen Pi	30.03	—
Jie Geng	30.09	—
Tian Hua Fen	30.48	—
Cang Zhu	30.67	—
Dang Shen	30.98	—
Zhi Ke	30.98	—
(Huai) Niu Xi	31.10	—
Ze Xie	31.19	—
(Bai) Dou Kou	31.47	—
Bai He	31.59	—
Hou Po	32.04	—
Jiao Gu Lan	32.06	—
Fang Feng	32.24	—
Du Zhong	32.96	—
Lian Qiao	33.33	—
Mao Dong Qing	33.59	—
Huang Bai	33.75	—
(Shi) Chang Pu	34.51	—
Xi Xian Cao	34.62	—
Zi Hua Di Ding	34.72	—
Dang Gui	34.72	—
Du Huo	34.76	—
Fu Zi	34.99	—
He Huan Pi	35.06	—
Gan Cao	35.54	—
Ban Xia (Jiang)	35.59	—
Dong Kui Zi	36.00	—
Shan Yu Rou	36.06	—
Hong Hua	36.10	—
Pu Gong Ying	36.57	—
Dan Dou Chi	36.62	—
Mu Gua	36.81	—
Zi Wan	36.81	—
Suan Zao Ren	36.82	—
Sang Ye	37.02	—
Mu Dan Pi	37.46	—
Dang Gui Wei	37.78	—
Bai Mao Gen	37.91	—
Lu Gen	38.46	—
Chi Shao (Yao)	38.65	—
Ye Jiao Teng	38.72	—
Bu Gu Zhi	38.80	—
Ren Dong Teng	39.12	—
Dan Shen	39.18	—
Huang Qi	39.62	—
Xin Yi	40.20	—
Hong Jing Tian	40.37	—
Zhe Bei Mu	40.80	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Yao	41.11	—
Nü Zhen Zi	41.34	—
Huang Lian	41.58	—
Tu Fu Ling	41.66	—
Yuan Zhi	41.86	—
Fu Ling	41.93	—
Bai Zhu	42.19	—
Mai Men Dong	42.28	—
Lian Zi	42.53	—
Mai Ya	42.61	—
Bai Xian Pi	43.64	—
Xian He Cao	44.19	—
Huang Bo	44.28	—
Da Zao	44.44	—
Yin Yang Huo	44.53	—
Yu Jin	44.72	—
Ling Zhi	44.74	—
San Leng	47.69	—
Xia Ku Cao	48.03	—
Qin Jiao	48.49	—
Hu Zhang	49.00	—
Zhu Ling	49.09	—

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	12.02
61898	61898	0.00	17.71
62155	62155	0.00	19.20
62755	62755	0.00	10.66
62156	62156	0.00	19.15
62756	62756	0.00	9.97
62504	62504	0.00	12.11
61897	61897	0.00	17.84

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 2.1-2020-01

Validated substance/substance group **Cang Zhu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60160-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G042H1402821	62387	40	from supplier
PhytoComm	Cang Zhu	G042H1402821	62388	40	from supplier
PhytoComm	Cang Zhu	G042HS301SG1	62653	40	from supplier
PhytoComm	Cang Zhu	G042HS301SG1	62654	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G042H1402821	62387 [†]	20
PhytoComm	Cang Zhu	G042H1402821	62388 [†]	20
PhytoComm	Cang Zhu	G042HS301SG1	62653 [†]	20
PhytoComm	Cang Zhu	G042HS301SG1	62654 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	0	160	0	25 641
Type B	0	80	0	13 087
Type C	1	0	1	851

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	100.0000 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.9419 % (> 99.3590 %)	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
Yuan Zhi	6.02	—
Shan Yao	7.07	—
She Gan	8.06	—
Chai Hu	8.15	—
Pi Pa Ye	8.42	—
Chen Pi	9.44	—
Lian Qiao	9.55	—
Gou Qi Zi	9.91	—
Dang Gui Wei	10.30	—
Chuan Xiong	11.09	—
Chuang Mu Xiang	12.28	—
Suan Zao Ren	12.29	—
Zhi Mu	12.33	—
Bai Bian Dou	12.46	—
Shen Qu	12.51	—
Bai Shao Yao	12.60	—
Jiao Gu Lan	12.85	—
Gua Lou	12.94	—
Ji Li	13.17	—
San Qi	13.18	—
Dang Shen	13.43	—
Jin Yin Hua	14.38	—
Bai Zi Ren	14.39	—
Bai Mao Gen	14.50	—
Du Huo	14.70	—
Zi Wan	14.83	—
Lian Zi	14.89	—
Mao Dong Qing	15.19	—
Dan Shen	15.60	—
Yu Zhu	15.67	—
(Huai) Niu Xi	16.97	—
Mu Dan Pi	17.75	—
Ren Dong Teng	17.80	—
Shan Yu Rou	17.84	—
Ban Xia (Jiang)	17.97	—
Hou Po	18.13	—
Ling Zhi	18.36	—
Zhe Bei Mu	18.38	—
Xiang Fu	18.40	—
Fang Feng	18.40	—
Hong Jing Tian	18.42	—
Huang Qi	18.71	—
Tai Zi Shen	18.96	—
Sang Zhi	19.20	—
Gan Jiang	19.20	—
Dang Gui	19.42	—
Lu Gen	19.81	—
Gui Zhi	19.95	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qing Pi	20.10	—
Xie Bai	20.47	—
Tian Hua Fen	20.52	—
Jie Geng	20.59	—
Fu Xiao Mai	20.83	—
Huo Ma Ren	20.83	—
(Shi) Chang Pu	21.02	—
Yi Mu Cao	21.09	—
San Leng	21.56	—
Fu Zi	21.59	—
Di Gu Pi	22.10	—
Ju Hua	22.35	—
Wu Wei Zi	23.06	—
Bai Zhu	23.13	—
Rou Gui	23.66	—
Sheng Jiang	23.73	—
He Huan Pi	24.04	—
Ci Wu Jia	24.12	—
Tu Fu Ling	24.27	—
Fu Ling	24.40	—
Long Yan Rou	24.88	—
Ye Jiao Teng	25.48	—
Zhi Ke	26.02	—
Long Dan (Cao)	26.07	—
Chi Shao (Yao)	26.20	—
Da Zao	26.22	—
Ba Ji Tian	26.30	—
Bai Xian Pi	26.32	—
Dan Dou Chi	26.62	—
Gou Teng	26.63	—
Fang Ji	27.16	—
Zhu Ru	27.98	—
Yin Yang Huo	28.39	—
Ji Xue Teng	28.81	—
Niu Bang Zi	29.03	—
Yan Hu Suo	29.08	—
Mang Xiao	29.26	—
Wu Yao	29.44	—
Mu Gua	29.62	—
Gan Cao	30.09	—
Lai Fu Zi	30.10	—
Tian Men Dong	30.24	—
Zi Su Zi	30.56	—
Tao Ren	30.73	—
Chuan Lian Zi	31.24	—
Jiang Huang	31.37	—
E Zhu	31.41	—
Ma Huang	31.69	—
Cang Er Zi	32.09	—
Fo Shou	32.45	—
Huang Qin	32.75	—
Yu Jin	32.96	—
Che Qian Zi	32.97	—
Ku Shen	32.97	—
Yi Yi Ren	33.27	—
Qiang Huo	33.50	—
Zi Hua Di Ding	33.69	—
Bai He	34.43	—
Ze Xie	35.74	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Du Zhong	36.41	—
Ma Huang Gen	37.71	—
Sha Ren	38.48	—
Wei Ling Xian	39.13	—
Wang Bu Liu Xing	39.26	—
(Bai) Dou Kou	40.55	—
Huang Lian	41.68	—
Xian He Cao	42.62	—
Huang Bo	42.78	—
Dong Gua Zi	42.83	—
Xin Yi	43.15	—
Ma Bian Cao	44.34	—
Pu Gong Ying	44.40	—
Zhi Zi	45.08	—
Mai Men Dong	45.60	—
Hong Hua	45.67	—
Sha Shen (Bei)	45.88	—
Gu Sui Bu	46.73	—
Huang Jing	46.78	—
Qin Jiao	47.00	—
Xu Duan	47.06	—
Fu Shen	48.13	—
Sang Ye	48.76	—
Sang Ji Shend	50.00	—

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	6.02
62654	62654	0.00	6.61
62387	62387	0.00	8.06
62388	62388	0.00	8.22

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 2.1-2020-01

Validated substance/substance group **Chai Hu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60010-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G055H1055821	62375	40	from supplier
PhytoComm	Chai Hu	G055H1055821	62376	40	from supplier
PhytoComm	Chai Hu	G055HS233SK1	62661	40	from supplier
PhytoComm	Chai Hu	G055HS233SK1	62662	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G055H1055821	62375 [†]	20
PhytoComm	Chai Hu	G055H1055821	62376 [†]	20
PhytoComm	Chai Hu	G055HS233SK1	62661 [†]	20
PhytoComm	Chai Hu	G055HS233SK1	62662 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G055H1055422	5
PhytoComm	Chai Hu	G055H1055322	1

- 845 spectra from 12 *Apo-Ident* customers from a total of 512 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	0	160	0	25 641
Type B	0	78	2	13 087
Type C	2	0	8	843

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	100.0000 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	97.5000 % (> 93.7500 %)
Type C	99.7907 % (> 99.2022 %)	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
Ji Li	5.58	—
Bai Bian Dou	6.12	—
Mao Dong Qing	7.52	—
Ling Zhi	7.81	—
Yuan Zhi	8.22	—
Hou Po	8.85	—
Gou Qi Zi	9.38	—
Shen Qu	9.39	—
Lian Zi	9.78	—
Ban Xia (Jiang)	9.95	—
Dan Dou Chi	10.10	—
Bai Xian Pi	10.37	—
Zhi Mu	10.94	—
Yu Jin	11.15	—
Dang Gui Wei	11.29	—
Lian Qiao	11.91	—
Du Huo	12.11	—
Cang Zhu	12.21	—
Fang Feng	12.38	—
Tu Fu Ling	12.65	—
Gua Lou	12.90	—
(Shi) Chang Pu	13.06	—
Qiang Huo	13.07	—
Ren Dong Teng	13.22	—
Yin Yang Huo	13.27	—
Zi Wan	13.57	—
Chuan Xiong	13.87	—
Pi Pa Ye	14.30	—
Fu Xiao Mai	14.57	—
She Gan	14.63	—
Jiao Gu Lan	14.70	—
Shan Yao	14.76	—
Gui Zhi	14.96	—
Lu Gen	14.98	—
Dang Gui	15.04	—
Bai Mao Gen	15.47	—
Bai Shao Yao	15.49	—
Che Qian Zi	15.58	—
(Huai) Niu Xi	15.65	—
Rou Gui	15.66	—
Dan Shen	16.09	—
Gan Jiang	16.21	—
Fu Zi	16.40	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qing Pi	16.62	—
Hong Jing Tian	17.08	—
Fu Ling	17.18	—
Sheng Jiang	17.60	—
Jin Yin Hua	17.96	—
Sang Zhi	18.40	—
Yan Hu Suo	18.45	—
Chen Pi	18.55	—
Yi Yi Ren	18.79	—
Ye Jiao Teng	19.37	—
Yi Mu Cao	19.72	—
Xie Bai	20.21	—
Di Gu Pi	20.50	—
Ju Hua	20.62	—
Ba Ji Tian	20.75	—
He Huan Pi	20.89	—
Du Zhong	21.13	—
San Leng	21.18	—
Tian Hua Fen	21.47	—
Huo Ma Ren	21.64	—
Xiang Fu	21.94	—
Fang Ji	22.08	—
Ma Huang Gen	22.65	—
Gou Teng	22.70	—
Bai Zi Ren	22.83	—
Mu Dan Pi	22.97	—
Huang Qi	23.46	—
Zhe Bei Mu	23.80	—
Wu Yao	23.95	—
Wu Wei Zi	24.33	—
Suan Zao Ren	24.57	—
Dang Shen	24.91	—
Jie Geng	25.41	—
Tai Zi Shen	25.87	—
Zhu Ru	25.91	—
Jiang Huang	25.99	—
Ma Huang	26.01	—
Tao Ren	26.22	—
Long Dan (Cao)	26.28	—
Zi Hua Di Ding	26.42	—
Ji Xue Teng	26.51	—
Xian He Cao	27.83	—
Mang Xiao	28.07	—
Wang Bu Liu Xing	28.13	—
Da Zao	28.50	—
Zi Su Zi	28.64	—
San Qi	28.68	—
Tian Men Dong	28.93	—
Ci Wu Jia	29.01	—
E Zhu	29.20	—
Huang Lian	29.47	—
Chuan Lian Zi	29.84	—
Long Yan Rou	29.85	—
Bai Zhu	30.93	—
Mu Gua	30.95	—
Lai Fu Zi	31.38	—
Niu Bang Zi	31.44	—
Yu Zhu	31.62	—
Ku Shen	31.75	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Bai	31.93	—
Sha Ren	32.43	—
Chuang Mu Xiang	32.52	—
Fo Shou	33.14	—
Zhi Ke	34.54	—
Shan Yu Rou	35.40	—
(Bai) Dou Kou	36.02	—
Ze Xie	37.09	—
Wei Ling Xian	37.38	—
Sha Shen (Bei)	37.53	—
Chi Shao (Yao)	38.23	—
Ma Bian Cao	38.52	—
Bai He	39.27	—
Qin Jiao	39.35	—
Huang Bo	39.35	—
Cang Er Zi	39.81	—
Huang Jing	40.29	—
Huang Qin	40.45	—
Xu Duan	40.47	—
Fu Shen	40.49	—
Gu Sui Bu	40.57	—
Pu Gong Ying	40.60	—
Xin Yi	41.10	—
Wu Zhu Yu	44.04	—
Nü Zhen Zi	44.10	—
Mai Men Dong	44.32	—
Sang Bai Pi	45.23	—
Xi Xian Cao	45.28	—
He Shou Wu	45.43	—
Zhi Zi	45.91	—
Ge Gen	46.15	—
Sang Ye	46.44	—
Yin Chen Hao	47.14	—
Dong Gua Zi	47.17	—
Sang Ji Shend	47.79	—
Mai Ya	47.88	—
Ban Zhi Lian	48.48	—
Gan Cao	48.51	—

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	6.12
62376	62376	0.00	9.53
62375	62375	0.00	9.38
62662	62662	0.00	5.58

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 2.1-2020-01

Validated substance/substance group **Che Qian Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60255-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Che Qian Zi	1	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

Validation samples

A total of 39 821 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 *Che Qian Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	G192H0723421	2
PhytoComm	Che Qian Zi	G192H0723321	2
PhytoComm	Che Qian Zi	G192H0723621	1

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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	79	1	25 721
Type B	0	40	0	13 127
Type C	0	0	6	847

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 % (> 99.9703 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8235 %)	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 *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	5.76	—
Yu Jin	7.96	—
Ling Zhi	9.68	—
Chai Hu	9.99	—
Yin Yang Huo	10.37	—
Ren Dong Teng	10.47	—
Bai Xian Pi	10.66	—
Bai Bian Dou	11.58	—
Ji Li	12.13	—
Du Zhong	12.18	—
Dan Dou Chi	12.26	—
Yan Hu Suo	12.35	—
Qiang Huo	13.73	—
Shen Qu	13.97	—
Fang Ji	14.03	—
Mao Dong Qing	14.46	—
Pi Pa Ye	14.50	—
Chuan Xiong	14.74	—
Yuan Zhi	15.11	—
Shan Yao	15.37	—
Ma Huang Gen	16.95	—
Ban Xia (Jiang)	17.91	—
Lian Zi	17.96	—
Lian Qiao	18.35	—
Dan Shen	19.71	—
He Huan Pi	19.95	—
Jiao Gu Lan	20.03	—
Qing Pi	20.09	—
Chen Pi	20.11	—
Ye Jiao Teng	20.28	—
Rou Gui	20.50	—
Yi Yi Ren	20.63	—
Tu Fu Ling	20.69	—
Hong Jing Tian	20.95	—
Gou Teng	21.05	—
Fu Xiao Mai	21.78	—
Sheng Jiang	22.45	—
Mu Dan Pi	22.75	—
Fu Ling	22.93	—
Ma Huang	23.17	—
Bai Shao Yao	23.19	—
Zi Hua Di Ding	23.19	—
Ji Xue Teng	23.45	—
Tian Hua Fen	23.46	—
Huo Ma Ren	23.52	—
Jin Yin Hua	24.38	—
Gui Zhi	24.61	—
Huang Lian	25.41	—
Gua Lou	25.70	—
Fu Zi	25.83	—
Zhu Ru	26.12	—
Tao Ren	27.47	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Wei Zi	27.79	—
Da Zao	28.74	—
Bai Zi Ren	29.63	—
Tai Zi Shen	29.96	—
Cang Zhu	30.02	—
Wang Bu Liu Xing	30.73	—
Di Gu Pi	31.65	—
Zhe Bei Mu	32.25	—
Huang Bai	33.14	—
Ci Wu Jia	33.29	—
Suan Zao Ren	33.97	—
Pu Gong Ying	34.07	—
Gou Qi Zi	34.76	—
Mang Xiao	35.37	—
Xi Xian Cao	35.88	—
Fo Shou	36.82	—
San Qi	37.18	—
Fu Shen	37.55	—
Zhi Ke	37.82	—
Tian Men Dong	40.47	—
Sang Zhi	41.04	—
Long Yan Rou	42.53	—
Jiang Huang	42.65	—
Jie Geng	42.77	—
Chi Shao (Yao)	43.21	—
Yu Zhu	43.21	—
Cang Er Zi	43.51	—
Shan Yu Rou	43.69	—
Ze Xie	44.21	—
Sang Ye	44.49	—
Dang Gui Wei	45.20	—
Xia Ku Cao	45.39	—
Huang Qin	45.59	—
Zhi Mu	45.77	—
Dang Gui	46.21	—
Bai He	47.35	—
Ban Zhi Lian	47.44	—
Chuang Mu Xiang	47.91	—
E Zhu	48.66	—
Zi Su Zi	50.31	—

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	6.10
62608	62608	0.00	5.76

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 2.1-2020-01

Validated substance/substance group **Chen Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60009-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G073HS254RR1	62445	40	from supplier
PhytoComm	Chen Pi	G073HS254RR1	62446	40	from supplier
PhytoComm	Chen Pi	G073HS254SG1	62669	40	from supplier
PhytoComm	Chen Pi	G073HS254SG1	62670	40	from supplier

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G073HS254RR1	62445 [†]	20
PhytoComm	Chen Pi	G073HS254RR1	62446 [†]	20
PhytoComm	Chen Pi	G073HS254SG1	62669 [†]	20
PhytoComm	Chen Pi	G073HS254SG1	62670 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	g07h1128221	1
Phytocomm	Chen Pi	g073h1128222	1
PhytoComm	Chen Pi	G073H1128222	1

- 849 spectra from 12 *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 *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	25 561
Type B	0	120	0	13 047
Type C	0	0	4	849

The substance/substance group *Chen 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 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
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 *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	6.98	—
Gou Qi Zi	7.50	—
Yu Zhu	7.63	—
Gua Lou	8.27	—
Cang Zhu	9.70	—
Fang Feng	9.75	—
Dang Gui Wei	10.08	—
Yuan Zhi	10.36	—
Zhi Mu	10.50	—
Fo Shou	10.90	—
Pi Pa Ye	11.07	—
Ye Jiao Teng	11.64	—
Bai Shao Yao	12.17	—
Da Zao	12.83	—
Wu Yao	13.76	—
Long Yan Rou	13.85	—
Ren Dong Teng	14.27	—
San Qi	14.35	—
Gu Sui Bu	14.41	—
Zhe Bei Mu	14.44	—
Shen Qu	14.77	—
Dang Shen	15.02	—
Xu Duan	15.03	—
Tian Men Dong	15.10	—
Gui Zhi	15.11	—
Hong Jing Tian	15.58	—
Shan Yao	15.77	—
Huang Jing	15.84	—
Lian Qiao	15.92	—
Tao Ren	15.92	—
Tai Zi Shen	16.10	—
Lian Zi	16.53	—
Lu Gen	16.72	—
Jin Yin Hua	16.73	—
Dan Shen	17.30	—
Tu Fu Ling	17.51	—
Ji Li	17.73	—
Yi Mu Cao	17.73	—
Bai Bian Dou	17.76	—
He Huan Pi	18.21	—
Bai Zi Ren	18.34	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang	18.51	—
Dang Gui	18.57	—
Yu Jin	19.17	—
Di Gu Pi	19.27	—
Fu Zi	19.28	—
Ma Bian Cao	19.38	—
Chuang Mu Xiang	19.75	—
Jiao Gu Lan	19.83	—
Rou Gui	20.21	—
Ci Wu Jia	20.37	—
Qiang Huo	20.66	—
Shan Yu Rou	20.85	—
Huo Ma Ren	20.86	—
Gou Teng	21.15	—
Bai Zhu	21.16	—
Jie Geng	21.51	—
Hou Po	21.89	—
Xian He Cao	21.93	—
Sang Bai Pi	22.12	—
Yin Chen Hao	22.51	—
Ju Hua	22.53	—
Mao Dong Qing	22.53	—
Ban Xia (Jiang)	22.79	—
Yin Yang Huo	23.01	—
Ji Xue Teng	23.26	—
Zhi Ke	23.33	—
Tian Hua Fen	23.57	—
Sha Ren	24.11	—
Jiang Huang	24.12	—
Bai Xian Pi	24.29	—
Wu Wei Zi	24.30	—
Sheng Jiang	24.57	—
Chai Hu	24.59	—
Fu Xiao Mai	24.72	—
Sang Ji Shend	25.06	—
Du Huo	25.08	—
Gan Cao	25.18	—
Fu Ling	25.24	—
Mu Dan Pi	25.43	—
Nü Zhen Zi	25.59	—
He Shou Wu	25.96	—
Ling Zhi	26.01	—
Wu Zhu Yu	26.49	—
Huang Qin	26.53	—
Gan Jiang	26.95	—
Huang Bo	27.66	—
Suan Zao Ren	27.79	—
Ze Xie	28.07	—
(Shi) Chang Pu	28.14	—
Dan Dou Chi	28.38	—
Yan Hu Suo	28.95	—
Sang Zhi	29.09	—
Mang Xiao	29.18	—
Zhu Ru	29.42	—
Che Qian Zi	29.50	—
Zi Wan	30.07	—
She Gan	30.41	—
Fang Ji	31.01	—
Qing Pi	31.42	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	32.26	—
Ba Ji Tian	33.40	—
Yi Yi Ren	33.52	—
Chi Shao (Yao)	33.62	—
Ma Huang Gen	33.83	—
Fu Shen	34.03	—
San Leng	34.45	—
Ge Gen	34.70	—
Xiang Fu	34.95	—
Xie Bai	35.33	—
Dan Zhu Ye	35.72	—
Huang Qi	36.18	—
Bu Gu Zhi	37.38	—
(Bai) Dou Kou	38.02	—
Hu Zhang	38.24	—
Mai Men Dong	38.38	—
Ku Shen	38.42	—
Cang Er Zi	38.49	—
Huang Lian	38.63	—
Zi Su Zi	39.07	—
Chuan Lian Zi	39.18	—
Bai Mao Gen	39.29	—
E Zhu	39.50	—
Zhi Shi	42.82	—
Xin Yi	42.97	—
Wei Ling Xian	43.00	—
Du Zhong	43.86	—
Niu Bang Zi	48.50	—
Wang Bu Liu Xing	49.04	—
(Huai) Niu Xi	50.00	—

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
62670	62670	0.00	6.98
62669	62669	0.00	7.18
62377	62377	0.00	9.75
62378	62378	0.00	10.96
62445	62445	0.00	7.63
62446	62446	0.00	8.12

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 2.1-2020-01

Validated substance/substance group **Chi Shao (Yao)**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60014-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Chi Shao (Yao)	2	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

Validation samples

A total of 39 821 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 *Chi Shao (Yao)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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)	G180H0705321	1
Phytocomm	Chi Shao (Yao)	g180h0705122	1
Phytocomm	Chi Shao (Yao)	G189H0705321	1
Phytocomm	Chi Shao (Yao)	G180H0705422	1

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	2	2	849

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8246 %)	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 *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	11.93	—
Bai Shao Yao	12.91	—
Ze Xie	13.61	—
She Gan	19.43	—
Dang Gui Wei	22.05	—
Zi Wan	22.13	—
Lian Qiao	24.41	—
Gou Qi Zi	24.76	—
Gan Cao	25.65	—
San Leng	26.02	—
(Huai) Niu Xi	26.18	—
Jie Geng	26.35	—
Yuan Zhi	26.78	—
Bai Mao Gen	27.43	—
Mu Gua	28.30	—
Lu Gen	29.56	—
Cang Zhu	29.58	—
Gui Zhi	30.14	—
Du Huo	31.55	—
Zhi Ke	32.72	—
Mu Dan Pi	32.77	—
Xiang Fu	32.92	—
Sang Zhi	33.56	—
Wu Yao	33.74	—
Chuan Xiong	34.43	—
(Shi) Chang Pu	34.58	—
Chen Pi	35.09	—
Di Gu Pi	35.19	—
Bai Zhu	35.98	—
Gua Lou	37.39	—
Chai Hu	38.01	—
Mang Xiao	38.10	—
Dang Gui	39.16	—
Huang Qin	39.23	—
Qing Pi	39.87	—
Ban Zhi Lian	40.26	—
Wei Ling Xian	41.25	—
Shan Yu Rou	41.48	—
Huang Lian	41.57	—
Ju Hua	42.73	—
Dan Shen	42.92	—
Hong Jing Tian	43.06	—
Long Dan (Cao)	43.27	—
Yi Mu Cao	43.33	—
Wu Wei Zi	44.28	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dang Shen	45.54	—
Fu Zi	45.61	—
Jin Yin Hua	45.83	—
Ma Huang	45.89	—
Shan Yao	46.29	—
Zi Hua Di Ding	46.64	—
Zhi Mu	46.72	—
Cang Er Zi	47.75	—
Niu Bang Zi	47.88	—
Qin Jiao	48.38	—
Da Zao	48.98	—
Yin Yang Huo	49.04	—
Tu Fu Ling	49.26	—
Ji Xue Teng	49.41	—
Long Yan Rou	49.78	—
Gan Jiang	49.80	—

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	12.54
62481	62481	0.00	13.61
62364	62364	0.00	11.93
62482	62482	0.00	13.75

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 2.1-2020-01

Validated substance/substance group	Chuan Lian Zi
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	12/03/2020
Report number	60078-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Chuan Lian Zi	1	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	G157H0306821	62355	40	from supplier
PhytoComm	Chuan Lian Zi	G157H0306821	62356	40	from supplier

Validation samples

A total of 39 821 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 Lian Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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	G157H0306821	62355 [†]	20
PhytoComm	Chuan Lian Zi	G157H0306821	62356 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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

- 850 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	3	0	3	847

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.7209 % (> 99.1338 %)	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 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
Bai He	9.15	—
San Qi	11.52	—
Sheng Jiang	12.49	—
Di Gu Pi	12.96	—
Sang Zhi	13.59	—
Mai Ya	15.69	—
Chuan Xiong	16.39	—
Jie Geng	16.59	—
Lai Fu Zi	17.10	—
E Zhu	18.29	—
Mai Men Dong	19.11	—
Tian Hua Fen	19.46	—
Long Dan (Cao)	20.38	—
Dang Gui	20.39	—
Dang Shen	21.14	—
Chai Hu	22.67	—
Gan Jiang	23.13	—
Bai Zhu	23.32	—
Gui Zhi	23.42	—
Ji Li	23.88	—
Mang Xiao	23.91	—
Ba Ji Tian	24.38	—
Qin Jiao	25.27	—
She Gan	25.40	—
Mu Dan Pi	25.69	—
Bai Bian Dou	26.10	—
Mo Yao	26.31	—
Cang Er Zi	27.05	—
Zhi Mu	27.23	—
Du Huo	27.30	—
Dang Gui Wei	27.44	—
Da Zao	27.59	—
Mu Gua	27.98	—
Yan Hu Suo	28.00	—
San Leng	28.18	—
Zi Su Zi	28.87	—
Ban Zhi Lian	28.98	—
Jiang Huang	29.02	—
Gou Qi Zi	29.65	—
Cang Zhu	32.25	—
Qing Pi	32.39	—
(Huai) Niu Xi	32.84	—
Huo Ma Ren	33.04	—
Xie Bai	33.10	—
Zhu Ling	33.12	—
Huang Qi	33.21	—
Lu Gen	33.39	—
Shan Yao	33.52	—
Sha Shen (Bei)	33.61	—
(Shi) Chang Pu	33.70	—
Ban Xia (Jiang)	34.07	—
Lian Zi	35.06	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xiang Fu	35.71	—
Gua Lou	36.23	—
Fang Feng	38.23	—
Fu Ling	38.99	—
Niu Bang Zi	39.45	—
Rou Gui	39.45	—
Tu Fu Ling	39.49	—
Chen Pi	39.67	—
Zi Wan	39.82	—
Sha Ren	40.43	—
(Bai) Dou Kou	40.56	—
He Huan Pi	42.48	—
Yi Yi Ren	42.71	—
Hou Po	44.01	—
Yi Mu Cao	44.64	—
Yuan Zhi	44.83	—
Zhu Ru	45.28	—
Bai Xian Pi	45.59	—
Bai Shao Yao	45.99	—
Wu Wei Zi	46.34	—
Wang Bu Liu Xing	46.37	—
Tian Men Dong	46.52	—
Ku Shen	46.92	—
Ren Dong Teng	47.42	—
Lian Qiao	47.53	—
Bai Mao Gen	47.67	—
Wu Yao	48.50	—
Qiang Huo	48.94	—
Ma Huang Gen	48.99	—
Xian He Cao	49.31	—
Chi Shao (Yao)	49.91	—
Gou Teng	50.13	—

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
62356	62356	0.00	11.34
62355	62355	0.00	9.15

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 2.1-2020-01

Validated substance/substance group **Chuan Xiong**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60013-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Chuan Xiong	4	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	G136H0305621	61975	40	from supplier
PhytoComm	Chuan Xiong	G136H0305621	61976	40	from supplier
PhytoComm	Chuan Xiong	G136H0305622	62291	40	from supplier
PhytoComm	Chuan Xiong	G136H0305622	62292	40	from supplier
PhytoComm	Chuan Xiong	G136HS031RR2	62457	40	from supplier
PhytoComm	Chuan Xiong	G136HS031RR2	62458	40	from supplier
PhytoComm	Chuan Xiong	G136HS031SH1	62689	40	from supplier
PhytoComm	Chuan Xiong	G136HS031SH1	62690	40	from supplier

Validation samples

A total of 39 821 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 *Chuan Xiong*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 25 481 spectra from a total of 315 batches from further 172 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 *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	G136H0305621	61975 [†]	20
PhytoComm	Chuan Xiong	G136H0305621	61976 [†]	20
PhytoComm	Chuan Xiong	G136H0305622	62291 [†]	20
PhytoComm	Chuan Xiong	G136H0305622	62292 [†]	20

continued on the next page

continued from previous page

Supplier	Substance	Batch	Sample ID	Spectra
PhytoComm	Chuan Xiong	G136HS031RR2	62457 [†]	20
PhytoComm	Chuan Xiong	G136HS031RR2	62458 [†]	20
PhytoComm	Chuan Xiong	G136HS031SH1	62689 [†]	20
PhytoComm	Chuan Xiong	G136HS031SH1	62690 [†]	20

- 13 007 spectra from a total of 316 batches from further 172 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	G136H0305221	1
Phytocomm	Chuan Xiong	g136h0305121	1
Phytocomm	Chuan Xiong	G136H0305421	2
PhytoComm	Chuan Xiong	G136H0305221	2

- 847 spectra from 12 *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 *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	320	0	25 481
Type B	0	158	2	13 007
Type C	4	1	5	843

The substance/substance group *Chuan Xiong* can be distinguished from other substances with some limitations. (*False-positive* classifications were obtained.) In order to make these figures comparable,

[†]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 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 % (> 98.1250 %)
Type B	100.0000 % (> 99.9421 %)	98.7500 % (> 96.8750 %)
Type C	99.7674 % (> 99.1792 %)	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 *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	4.57	—
Shan Yao	5.76	—
Shen Qu	6.59	—
San Qi	6.87	—
Dang Gui	7.90	—
Long Dan (Cao)	7.92	—
Dang Shen	7.99	—
Cang Zhu	8.83	—
Ren Dong Teng	8.85	—
Yuan Zhi	9.12	—
Lian Qiao	9.40	—
Shan Yu Rou	9.50	—
Hong Jing Tian	9.65	—
Jin Yin Hua	9.78	—
Bai Bian Dou	10.49	—
Dan Shen	10.55	—
Chen Pi	10.71	—
Sang Zhi	11.41	—
Hou Po	11.57	—
Ji Li	11.59	—
Dang Gui Wei	11.68	—
Gou Qi Zi	11.99	—
Gua Lou	12.18	—
Lian Zi	12.25	—
Bai Shao Yao	12.37	—
Zhe Bei Mu	12.45	—
Fu Zi	12.54	—
San Leng	12.63	—
Mu Dan Pi	13.40	—
Mao Dong Qing	13.61	—
Ling Zhi	13.63	—
Di Gu Pi	13.78	—
Jiao Gu Lan	14.57	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jie Geng	14.65	—
Tian Hua Fen	15.19	—
Tai Zi Shen	15.23	—
Zhi Mu	15.76	—
Yu Zhu	15.93	—
Chai Hu	16.17	—
Bai Zi Ren	16.36	—
Dan Dou Chi	16.49	—
Huo Ma Ren	16.88	—
Ye Jiao Teng	17.28	—
Yin Yang Huo	17.49	—
He Huan Pi	17.54	—
Ba Ji Tian	17.58	—
She Gan	17.77	—
Ban Xia (Jiang)	18.16	—
Bai Xian Pi	18.23	—
Tu Fu Ling	18.39	—
Gui Zhi	18.84	—
Fu Xiao Mai	18.95	—
Chuang Mu Xiang	19.07	—
Gou Teng	19.35	—
Sheng Jiang	19.59	—
Du Huo	19.85	—
Chuan Lian Zi	19.93	—
Ma Huang	19.99	—
Ci Wu Jia	20.13	—
Fu Ling	20.29	—
Yan Hu Suo	20.53	—
Da Zao	20.71	—
Che Qian Zi	21.11	—
Ji Xue Teng	21.13	—
Tao Ren	21.18	—
Zhi Ke	21.26	—
Long Yan Rou	21.39	—
Wu Wei Zi	21.52	—
(Huai) Niu Xi	21.99	—
Bai He	22.20	—
Yu Jin	22.35	—
Rou Gui	22.48	—
Qing Pi	22.51	—
E Zhu	22.61	—
Fang Ji	23.24	—
Bai Mao Gen	23.30	—
Huang Qi	23.78	—
Qiang Huo	23.97	—
Fo Shou	24.02	—
Bai Zhu	24.19	—
Suan Zao Ren	24.20	—
Lu Gen	24.30	—
Zhu Ru	24.41	—
Ma Huang Gen	25.16	—
Sha Shen (Bei)	25.37	—
Zi Hua Di Ding	25.45	—
Gan Jiang	25.64	—
Mang Xiao	25.71	—
Fang Feng	26.17	—
Huang Qin	26.68	—
Gan Cao	26.70	—
Yi Yi Ren	27.56	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qin Jiao	27.61	—
Mu Gua	27.64	—
Tian Men Dong	28.36	—
Mai Ya	28.94	—
Lai Fu Zi	29.10	—
Du Zhong	32.22	—
Mai Men Dong	32.45	—
Huang Lian	32.54	—
Zi Wan	32.67	—
Ze Xie	33.96	—
Xian He Cao	34.22	—
Xie Bai	35.38	—
(Shi) Chang Pu	36.32	—
Cang Er Zi	36.34	—
Chi Shao (Yao)	36.84	—
Wu Yao	37.11	—
Jiang Huang	37.79	—
Zhu Ling	38.26	—
Fu Shen	38.33	—
Zi Su Zi	39.65	—
Ban Zhi Lian	41.60	—
Wang Bu Liu Xing	42.14	—
Sha Ren	42.57	—
Huang Bai	42.58	—
Yi Mu Cao	42.58	—
Hong Hua	42.79	—
Ju Hua	43.08	—
Xiang Fu	43.37	—
Mo Yao	46.50	—
Sang Ye	47.81	—
(Bai) Dou Kou	47.98	—
Pu Gong Ying	48.32	—
Niu Bang Zi	49.40	—

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
61975	61975	0.00	8.05
62689	62689	0.00	5.24
62690	62690	0.00	4.57
62458	62458	0.00	6.87
62292	62292	0.00	7.90
61976	61976	0.00	8.48
62457	62457	0.00	7.50
62291	62291	0.00	7.92

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 2.1-2020-01

Validated substance/substance group **Chuang Mu Xiang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50323-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Chuang Mu Xiang; Vladimira radix

Special notes

When selecting the *Chuang Mu Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Chuang Mu Xiang	1	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

Validation samples

A total of 39 821 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 *Chuang Mu Xiang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	G219H0309322	1
Phytocomm	Chuang Mu Xiang	g219h0309321	1
Phytocomm	Chuang Mu Xiang	g219h0309322	1
Phytocomm	Chuang Mu Xiang	G219H0309521	2

- 848 spectra from 12 *Apo-Ident* customers from a total of 512 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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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 *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
Chen Pi	9.92	—
Chuan Xiong	12.99	—
Zhi Mu	13.08	—
Gou Qi Zi	13.51	—
Yuan Zhi	14.08	—
Dang Shen	14.49	—
Cang Zhu	14.76	—
Shan Yao	15.91	—
Mu Dan Pi	16.44	—
Pi Pa Ye	16.59	—
Shan Yu Rou	17.26	—
Jin Yin Hua	18.06	—
Long Yan Rou	18.19	—
Hong Jing Tian	18.20	—
Bai Shao Yao	18.37	—
Yu Zhu	18.48	—
Bai Zhu	18.81	—
Suan Zao Ren	18.89	—
Jiao Gu Lan	19.85	—
Dang Gui Wei	20.08	—
Lian Qiao	20.24	—
Dan Shen	20.26	—
Lian Zi	20.31	—
Zhe Bei Mu	22.24	—
Bai Bian Dou	22.24	—
Shen Qu	22.52	—
Di Gu Pi	22.80	—
Ji Li	22.87	—
San Qi	22.93	—
Gua Lou	23.26	—
Tai Zi Shen	24.05	—
Ren Dong Teng	24.71	—
Chai Hu	25.31	—
Jie Geng	25.71	—
Mao Dong Qing	25.72	—
Gan Cao	25.97	—
Da Zao	26.34	—
Zhi Ke	26.64	—
Ling Zhi	27.02	—
Hou Po	27.07	—
Wu Wei Zi	27.32	—
Bai Zi Ren	27.45	—
Ci Wu Jia	28.17	—
Dang Gui	28.63	—
Tu Fu Ling	29.03	—
Ze Xie	29.24	—
Bai Xian Pi	29.92	—
Gui Zhi	30.03	—
Ban Xia (Jiang)	30.30	—
He Huan Pi	30.43	—
Zi Hua Di Ding	30.44	—
Cang Er Zi	30.84	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Zhi	31.82	—
Qing Pi	31.87	—
Sheng Jiang	32.03	—
Ye Jiao Teng	32.08	—
Fu Zi	32.23	—
Tian Men Dong	32.27	—
Huo Ma Ren	32.41	—
Yin Yang Huo	32.77	—
Tao Ren	32.83	—
Ji Xue Teng	32.96	—
Gou Teng	33.14	—
Tian Hua Fen	33.45	—
Dan Dou Chi	33.48	—
Fu Xiao Mai	33.89	—
Rou Gui	34.15	—
Huang Qin	34.83	—
Yan Hu Suo	35.74	—
Fo Shou	35.82	—
Fang Ji	35.82	—
Fu Ling	35.99	—
Mai Men Dong	36.59	—
San Leng	36.59	—
Ma Huang	36.94	—
Mang Xiao	39.06	—
Chi Shao (Yao)	39.50	—
Chuan Lian Zi	40.51	—
Huang Qi	40.95	—
Zhu Ru	41.06	—
Qiang Huo	41.61	—
Yu Jin	41.89	—
Hong Hua	41.93	—
Che Qian Zi	42.03	—
Zi Wan	42.22	—
Niu Bang Zi	43.32	—
Yi Yi Ren	43.48	—
Du Zhong	44.36	—
Yi Mu Cao	45.83	—
Huang Lian	46.38	—
Zi Su Zi	47.21	—
Ma Huang Gen	47.48	—
Bai Mao Gen	47.63	—
(Shi) Chang Pu	47.74	—
Huang Bai	48.93	—
Lai Fu Zi	49.79	—

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	9.92
62644	62644	0.00	9.94

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 2.1-2020-01

Validated substance/substance group **Ci Wu Jia**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60354-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	0	76	4	25 721
Type B	0	33	7	13 127
Type C	0	0	0	853

The substance/substance group *Ci Wu Jia* 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 % (> 99.9703 %)	95.0000 % (> 91.2500 %)
Type B	100.0000 % (> 99.9424 %)	82.5000 % (> 75.0000 %)
Type C	100.0000 % (> 98.8323 %)	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	2.92	—
He Huan Pi	3.48	—
Ban Xia (Jiang)	4.33	—
Rou Gui	4.66	—
Gui Zhi	5.18	—
Tao Ren	5.29	—
Fu Xiao Mai	5.71	—
Zhu Ru	6.13	—
Di Gu Pi	7.96	—
Fu Ling	8.47	—
Yi Yi Ren	8.87	—
Bai Bian Dou	9.28	—
Gou Teng	9.38	—
Ji Li	9.48	—
Bai Zi Ren	9.63	—
Huo Ma Ren	10.23	—
Lian Zi	10.56	—
Ling Zhi	11.96	—
Bai Shao Yao	12.11	—
Tai Zi Shen	12.56	—
Zhe Bei Mu	12.65	—
Shen Qu	14.12	—
Bai Xian Pi	14.12	—
Wang Bu Liu Xing	15.41	—
Tian Hua Fen	15.56	—
Ma Huang Gen	16.06	—
Ji Xue Teng	16.44	—
Fo Shou	18.73	—
Gua Lou	19.02	—
Ye Jiao Teng	19.25	—
Tian Men Dong	20.09	—
Ren Dong Teng	21.07	—
Chen Pi	21.38	—
Fu Shen	21.94	—
Yu Jin	22.07	—
Chuan Xiong	22.30	—
Yuan Zhi	22.60	—
Shan Yao	23.24	—
Long Yan Rou	23.36	—
Lian Qiao	23.84	—
Pi Pa Ye	24.04	—
Tu Fu Ling	24.16	—
Da Zao	24.53	—
Yan Hu Suo	25.28	—
Dan Shen	25.86	—
Dan Dou Chi	26.81	—
Mao Dong Qing	27.15	—
Dang Shen	30.17	—
Fu Zi	30.49	—
Zhi Mu	31.77	—
Jiao Gu Lan	31.91	—
Yin Yang Huo	32.39	—
Dang Gui Wei	33.11	—
Shan Yu Rou	33.80	—
Ma Huang	33.83	—
Yu Zhu	33.96	—
Fang Ji	33.97	—
San Qi	34.00	—
Chai Hu	34.35	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	34.85	—
Hong Jing Tian	35.39	—
Wu Wei Zi	35.69	—
Jin Yin Hua	37.50	—
Cang Zhu	37.65	—
Mang Xiao	38.11	—
Dang Gui	38.13	—
Che Qian Zi	38.35	—
Hou Po	38.83	—
Huang Qin	39.44	—
Qiang Huo	44.83	—
Huang Lian	46.14	—
Suan Zao Ren	46.58	—
Bai Zhu	47.13	—
Sang Zhi	47.47	—
Gou Qi Zi	48.23	—
Jie Geng	48.52	—
Ze Xie	48.84	—
Qing Pi	49.00	—

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
62624	62624	0.00	2.92
62623	62623	0.00	3.19

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 2.1-2020-01

Validated substance/substance group **Da Huang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60140-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Da Huang	2	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	G212H0320622	62129	40	from supplier
PhytoComm	Da Huang	G212H0320622	62130	40	from supplier
PhytoComm	Da Huang	G212H0320622	62157	40	from supplier
PhytoComm	Da Huang	G212H0320622	62158	40	from supplier
PhytoComm	Da Huang	G212H0320822	62495	40	from supplier
PhytoComm	Da Huang	G212H0320822	62496	40	from supplier

Validation samples

A total of 39 821 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 *Da Huang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 561 spectra from a total of 317 batches from further 172 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 *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	G212H0320622	62129 [†]	20
PhytoComm	Da Huang	G212H0320622	62130 [†]	20
PhytoComm	Da Huang	G212H0320622	62157 [†]	20
PhytoComm	Da Huang	G212H0320622	62158 [†]	20
PhytoComm	Da Huang	G212H0320822	62495 [†]	20
PhytoComm	Da Huang	G212H0320822	62496 [†]	20

- 13 047 spectra from a total of 318 batches from further 172 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	G212H0320321	2
Phytocomm	Da Huang	g212h0320023	1
Phytocomm	Da Huang	G212H0320521	3
PhytoComm	Da Huang	G212H0320321	1

- 846 spectra from 12 *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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	5	2	846

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 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8232 %)	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 *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	33.89	—

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
62158	62158	0.00	33.96
62495	62495	0.00	35.68
62130	62130	0.00	35.79
62157	62157	0.00	33.89
62496	62496	0.00	35.65
62129	62129	0.00	35.79

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 2.1-2020-01

Validated substance/substance group **Da Zao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60055-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Da Zao	3	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	G129H0316621	62215	40	from supplier
PhytoComm	Da Zao	G129H0316621	62216	40	from supplier
PhytoComm	Da Zao	G129HS019RT1	62455	40	from supplier
PhytoComm	Da Zao	G129HS019RT1	62456	40	from supplier
PhytoComm	Da Zao	G129H0316821	62539	40	from supplier
PhytoComm	Da Zao	G129H0316821	62540	40	from supplier

Validation samples

A total of 39 821 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 *Da Zao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G129H0316621	62215 [†]	20
PhytoComm	Da Zao	G129H0316621	62216 [†]	20
PhytoComm	Da Zao	G129HS019RT1	62455 [†]	20
PhytoComm	Da Zao	G129HS019RT1	62456 [†]	20
PhytoComm	Da Zao	G129H0316821	62539 [†]	20
PhytoComm	Da Zao	G129H0316821	62540 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 849 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	4	0	849

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8246 %)	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 *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
Gua Lou	13.97	—
Tian Men Dong	14.31	—
Long Yan Rou	14.49	—
Zhi Mu	15.42	—
Rou Gui	17.01	—
Dang Gui	18.41	—
Dang Shen	18.69	—
Bai Shao Yao	19.36	—
Chen Pi	19.43	—
Di Gu Pi	19.53	—
San Leng	20.45	—
Fu Xiao Mai	20.98	—
Jie Geng	21.42	—
Yu Zhu	22.00	—
Tai Zi Shen	22.16	—
Mai Men Dong	22.16	—
Ye Jiao Teng	22.55	—
Yuan Zhi	22.99	—
Chuan Xiong	23.41	—
He Huan Pi	23.50	—
Tao Ren	23.73	—
Fu Shen	23.90	—
Fo Shou	23.90	—
Lian Zi	24.05	—
Ban Xia (Jiang)	24.14	—
Ji Li	24.23	—
Shen Qu	24.27	—
Gou Teng	24.29	—
Gui Zhi	24.32	—
Gou Qi Zi	25.20	—
She Gan	25.57	—
Huang Qi	25.65	—
Long Dan (Cao)	25.96	—
Mu Dan Pi	26.04	—
Mang Xiao	26.05	—
Pi Pa Ye	26.19	—
Sheng Jiang	26.51	—
Dang Gui Wei	26.60	—
Bai Zhu	26.63	—
Fu Ling	26.75	—
Du Huo	26.84	—
Bai He	26.92	—
Sang Zhi	27.10	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qin Jiao	27.11	—
Bai Xian Pi	28.19	—
Zhe Bei Mu	28.25	—
Ji Xue Teng	28.56	—
Ma Huang	28.76	—
Huang Qin	29.22	—
Cang Zhu	29.63	—
Yi Yi Ren	30.28	—
Shan Yu Rou	30.61	—
Tu Fu Ling	30.64	—
Bai Bian Dou	30.94	—
Ci Wu Jia	32.70	—
Ma Huang Gen	32.99	—
Mu Gua	33.54	—
Lian Qiao	33.64	—
Bai Zi Ren	33.68	—
San Qi	34.05	—
Zhu Ru	34.39	—
Bai Mao Gen	34.86	—
Mai Ya	34.89	—
Zhi Ke	35.49	—
Ling Zhi	35.95	—
Chai Hu	35.97	—
(Huai) Niu Xi	36.37	—
Dan Shen	37.05	—
Huo Ma Ren	37.24	—
Yin Yang Huo	37.74	—
Ba Ji Tian	37.96	—
Ze Xie	38.08	—
Lai Fu Zi	38.10	—
Shan Yao	38.19	—
Ren Dong Teng	38.34	—
Jiao Gu Lan	38.57	—
Chuang Mu Xiang	39.36	—
Gan Jiang	39.86	—
Yan Hu Suo	40.83	—
Chuan Lian Zi	41.81	—
Sha Shen (Bei)	42.34	—
Lu Gen	42.36	—
Zi Wan	42.53	—
Tian Hua Fen	42.78	—
Wu Wei Zi	43.62	—
Hong Jing Tian	44.28	—
E Zhu	44.29	—
Gan Cao	45.26	—
Fu Zi	45.43	—
Mao Dong Qing	45.60	—
Fang Feng	46.08	—
Chi Shao (Yao)	46.71	—
Dan Dou Chi	46.91	—
Xian He Cao	47.37	—
Hou Po	47.71	—
Jin Yin Hua	47.76	—
Fang Ji	48.44	—
Yu Jin	48.65	—
Wang Bu Liu Xing	49.58	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/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
62215	62215	0.00	20.45
62456	62456	0.00	13.97
62539	62539	0.00	18.77
62540	62540	0.00	18.41
62455	62455	0.00	14.14
62216	62216	0.00	21.46

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 2.1-2020-01

Validated substance/substance group **Dan Dou Chi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60794-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dan Dou Chi	1	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

Validation samples

A total of 39 821 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 Dou Chi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	4	76	4	25 717
Type B	0	37	3	13 127
Type C	0	0	1	852

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	99.9855 % (> 99.9706 %)	95.0000 % (> 91.2500 %)
Type B	100.0000 % (> 99.9424 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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 *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	3.55	—
Ling Zhi	6.57	—
Bai Bian Dou	7.67	—
Ye Jiao Teng	8.28	—
Lian Zi	8.42	—
Mao Dong Qing	9.11	—
Ren Dong Teng	9.15	—
Tu Fu Ling	9.45	—
Yuan Zhi	11.10	—
Yin Yang Huo	11.21	—
Hou Po	11.36	—
Yu Jin	11.93	—
Chai Hu	11.96	—
Rou Gui	13.13	—
Shen Qu	14.17	—
Ban Xia (Jiang)	14.70	—
Qiang Huo	16.45	—
Ji Li	16.73	—
Che Qian Zi	17.03	—
Sheng Jiang	17.25	—
Gou Teng	17.49	—
He Huan Pi	17.55	—
Ma Huang Gen	17.83	—
Gui Zhi	18.21	—
Dan Shen	18.40	—
Yan Hu Suo	18.47	—
Ji Xue Teng	18.60	—
Fu Xiao Mai	19.29	—
Yi Yi Ren	19.44	—
Pi Pa Ye	19.64	—
Shan Yao	19.67	—
Lian Qiao	19.68	—
Huo Ma Ren	19.81	—
Chuan Xiong	20.35	—
Jin Yin Hua	21.18	—
Jiao Gu Lan	21.51	—
Ma Huang	21.54	—
Chen Pi	21.76	—
Bai Shao Yao	22.12	—
Fu Ling	22.87	—
Gua Lou	23.17	—
Tao Ren	23.31	—
Fang Ji	23.74	—
Qing Pi	24.02	—
Du Zhong	24.28	—
Fu Zi	24.43	—
Tian Hua Fen	24.68	—
Zhe Bei Mu	24.87	—
Hong Jing Tian	24.97	—
Mu Dan Pi	26.16	—
Di Gu Pi	26.87	—
Cang Zhu	27.40	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	27.40	—
Zi Hua Di Ding	27.88	—
Long Yan Rou	28.37	—
Tai Zi Shen	29.27	—
Ci Wu Jia	29.35	—
Huang Lian	29.80	—
Wu Wei Zi	30.17	—
Gou Qi Zi	30.59	—
Bai Zi Ren	30.89	—
Fu Shen	31.96	—
Tian Men Dong	32.18	—
Fo Shou	32.74	—
Da Zao	32.77	—
Zhu Ru	33.91	—
Yu Zhu	35.06	—
Mang Xiao	35.61	—
Suan Zao Ren	36.90	—
Wang Bu Liu Xing	37.24	—
San Qi	37.34	—
Dang Gui Wei	38.09	—
Sang Zhi	38.29	—
Ban Zhi Lian	38.39	—
Huang Bai	39.02	—
Chuang Mu Xiang	41.14	—
Jie Geng	41.33	—
Dang Shen	41.42	—
Zhi Mu	41.66	—
Dang Gui	42.08	—
Shan Yu Rou	42.54	—
Huang Qin	42.72	—
Ze Xie	42.78	—
Chi Shao (Yao)	42.94	—
Pu Gong Ying	43.71	—
Bai He	46.43	—
Chuan Lian Zi	46.47	—
Bai Zhu	46.82	—
Zi Su Zi	47.19	—
Cang Er Zi	50.32	—

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	3.55
62714	62714	0.00	4.15

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 2.1-2020-01

Validated substance/substance group **Dan Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60030-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dan Shen	2	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	G214H0441821	62309	40	from supplier
PhytoComm	Dan Shen	G214H0441821	62310	40	from supplier
PhytoComm	Dan Shen	G214HS059SG1	62573	40	from supplier
PhytoComm	Dan Shen	G214HS059SG1	62574	40	from supplier

Validation samples

A total of 39 821 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 Shen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G214H0441821	62309 [†]	20
PhytoComm	Dan Shen	G214H0441821	62310 [†]	20
PhytoComm	Dan Shen	G214HS059SG1	62573 [†]	20
PhytoComm	Dan Shen	G214HS059SG1	62574 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G214H0441221	1
Phytocomm	Dan Shen	G214H0441322	1
Phytocomm	Dan Shen	G214H0441522	1
Phytocomm	Dan Shen	g214h0441121	1
PhytoComm	Dan Shen	G214H0441221	2
PhytoComm	Dan Shen	G214H0441322	1
PhytoComm	Dan Shen	G214H0441522	1

- 844 spectra from 12 *Apo-Ident* customers from a total of 511 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	159	1	25 641
Type B	5	79	1	13 082
Type C	2	0	9	842

The substance/substance group *Dan 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 %)	99.3750 % (> 97.5000 %)
Type B	99.9564 % (> 99.9275 %)	98.7500 % (> 95.0000 %)
Type C	99.5349 % (> 98.9463 %)	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 *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	4.47	—
Mu Dan Pi	6.37	—
Hou Po	10.51	—
Qiang Huo	11.01	—
Jin Yin Hua	11.03	—
Hong Jing Tian	11.50	—
Yin Yang Huo	11.77	—
Ye Jiao Teng	11.95	—
Zhi Ke	12.03	—
Jiao Gu Lan	13.41	—
Gu Sui Bu	13.85	—
Huang Bai	14.27	—
Xu Duan	14.47	—
Ma Huang	16.14	—
Fang Ji	16.15	—
Chai Hu	16.27	—
Chuan Xiong	16.36	—
Bai Xian Pi	16.60	—
Ji Xue Teng	16.81	—
Ma Bian Cao	16.83	—
He Huan Pi	16.89	—
Sang Ji Shend	17.00	—
Yi Mu Cao	17.13	—
Shan Yao	17.17	—
Yin Chen Hao	17.32	—
Pi Pa Ye	17.63	—
Dan Dou Chi	17.93	—
Wu Zhu Yu	18.03	—
Ren Dong Teng	18.16	—
Nü Zhen Zi	18.40	—
Qing Pi	18.83	—
Huang Lian	18.85	—
Mao Dong Qing	19.25	—
Bai Shao Yao	19.63	—
Wu Wei Zi	19.64	—
Huang Bo	19.91	—
Fu Zi	19.94	—
Yuan Zhi	20.08	—
Yu Jin	20.82	—
Bai Bian Dou	21.23	—
Du Zhong	21.25	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Ren	21.36	—
Lian Zi	21.61	—
Xian He Cao	21.75	—
Di Gu Pi	21.87	—
Gui Zhi	22.01	—
Gou Qi Zi	22.25	—
Yan Hu Suo	22.31	—
Dan Zhu Ye	22.43	—
Gua Lou	22.95	—
Lian Qiao	23.96	—
Che Qian Zi	24.56	—
He Shou Wu	24.58	—
Tu Fu Ling	24.98	—
Sang Bai Pi	25.35	—
Wu Yao	25.50	—
Shen Qu	25.90	—
Ling Zhi	26.11	—
Cang Zhu	26.19	—
Ji Li	26.70	—
Chen Pi	27.01	—
Hu Zhang	27.61	—
Pu Gong Ying	27.97	—
Ju Hua	28.09	—
Fang Feng	28.10	—
Rou Gui	28.29	—
Bu Gu Zhi	28.42	—
Shan Yu Rou	28.62	—
Gou Teng	29.09	—
Tao Ren	30.72	—
Chi Shao (Yao)	30.88	—
Bai Zhu	30.91	—
Wei Ling Xian	31.18	—
Cang Er Zi	31.95	—
Sang Ye	32.18	—
Da Zao	32.52	—
Ze Xie	32.62	—
Zhe Bei Mu	32.65	—
Ban Zhi Lian	32.75	—
Huang Qin	32.84	—
Ge Gen	33.20	—
Huo Ma Ren	33.28	—
Gan Cao	34.22	—
Jing Jie	34.53	—
Xiang Fu	34.60	—
Jie Geng	34.74	—
Jiang Huang	35.24	—
Chuang Mu Xiang	35.39	—
Bai Mao Gen	35.54	—
San Qi	36.08	—
Xia Ku Cao	36.24	—
Sheng Jiang	36.52	—
Sang Zhi	36.56	—
Tian Hua Fen	36.65	—
Mang Xiao	36.94	—
Tai Zi Shen	37.01	—
Zhi Mu	37.19	—
Ma Huang Gen	37.76	—
Dang Gui	38.22	—
San Leng	38.30	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Shi	38.69	—
Long Yan Rou	39.23	—
Hong Hua	39.27	—
Dang Gui Wei	39.37	—
(Shi) Chang Pu	39.83	—
Ban Xia (Jiang)	40.39	—
Yu Zhu	40.44	—
Lu Gen	40.68	—
Fu Xiao Mai	40.73	—
Xi Xian Cao	40.79	—
Ci Wu Jia	41.15	—
Suan Zao Ren	41.82	—
Dang Shen	42.37	—
Huang Jing	43.46	—
Fu Ling	43.49	—
Ku Shen	43.92	—
Bai Zi Ren	44.20	—
Chuan Lian Zi	45.04	—
Fu Shen	45.07	—
Zi Wan	45.89	—
Yi Yi Ren	46.04	—
Xin Yi	46.26	—
Bai Hua She She Cao	46.57	—
Long Dan (Cao)	46.82	—
Han Lian Cao	47.21	—
(Bai) Dou Kou	47.36	—
E Zhu	47.59	—
Bai He	48.67	—
Wu Mei	49.61	—
She Gan	49.72	—
Fo Shou	50.00	—

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	4.47
62309	62309	0.00	11.13
62310	62310	0.00	11.01
62574	62574	0.00	8.19

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 2.1-2020-01

Validated substance/substance group **Dan Zhu Ye**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60147-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G046H1105621	62105	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105621	62106	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105821	62581	40	from supplier
PhytoComm	Dan Zhu Ye	G046H1105821	62582	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G046H1105621	62105 [†]	20
PhytoComm	Dan Zhu Ye	G046H1105621	62106 [†]	20
PhytoComm	Dan Zhu Ye	G046H1105821	62581 [†]	20
PhytoComm	Dan Zhu Ye	G046H1105821	62582 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	158	2	25 641
Type B	1	78	2	13 086
Type C	14	0	2	837

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 % (> 99.9701 %)	98.7500 % (> 96.8750 %)
Type B	99.9855 % (> 99.9566 %)	97.5000 % (> 93.7500 %)
Type C	98.1279 % (> 97.5418 %)	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
Ma Bian Cao	5.79	—
Yin Chen Hao	9.40	—
Bu Gu Zhi	9.53	—
He Shou Wu	10.65	—
Jing Jie	10.80	—
Xu Duan	12.15	—
Sang Bai Pi	12.57	—
Sha Ren	12.74	—
Gu Sui Bu	16.69	—
Du Zhong	17.14	—
Xian He Cao	18.00	—
Sang Ji Shend	20.47	—
Gui Zhi	20.90	—
He Huan Pi	22.00	—
Wu Zhu Yu	22.93	—
Nü Zhen Zi	22.97	—
Yi Mu Cao	23.25	—
Yin Yang Huo	23.41	—
Dan Shen	24.94	—
Gou Teng	26.22	—
Jin Qian Cao	26.75	—
Qiang Huo	27.58	—
(Sheng) Di Huang	29.01	—
Yu Jin	29.16	—
Chen Pi	29.89	—
Pu Gong Ying	31.79	—
Huang Bo	32.21	—
Wu Jia Pi	32.67	—
Fu Zi	32.71	—
Zhi Shi	33.94	—
Hou Po	34.98	—
Yu Xing Cao	36.30	—
Fang Feng	36.42	—
Ji Li	36.95	—
Bai Hua She She Cao	37.76	—
Hong Jing Tian	37.80	—
Jiang Huang	38.05	—
Ge Gen	38.15	—
Mang Xiao	39.52	—
Huang Lian	39.89	—
Sang Ye	39.90	—
Xuan Shen	40.00	—
Hu Zhang	40.62	—
Xin Yi	41.29	—
Xia Ku Cao	41.42	—
(Bai) Dou Kou	42.72	—
Wu Yao	43.51	—
Han Lian Cao	44.60	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wei Ling Xian	46.60	—
Tu Si Zi	47.44	—
Hua Shi	48.87	—
Lu Gen	49.69	—

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
62582	62582	0.00	10.24
62106	62106	0.00	6.83
62105	62105	0.00	5.79
62581	62581	0.00	10.11

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 2.1-2020-01

Validated substance/substance group **Dang Gui**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60003-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dang Gui	2	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

Validation samples

A total of 39 821 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 *Dang Gui*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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	G022H1306422	3
PhytoComm	Dang Gui	G13-0784	1
PhytoComm	Dang Gui	G022H1306121	1
PhytoComm	Dang Gui	G022H1306621	4

- 841 spectra from 12 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	80	0	13 087
Type C	1	1	11	840

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.9225 % (> 99.3337 %)	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 *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
Chuan Xiong	5.06	—
Long Dan (Cao)	7.00	—
Dang Gui Wei	7.32	—
Gou Qi Zi	8.84	—
Dang Shen	8.93	—
San Leng	10.72	—
Sang Zhi	11.24	—
Zhi Mu	11.71	—
Cang Zhu	11.80	—
Yuan Zhi	11.82	—
San Qi	12.13	—
Shan Yu Rou	12.15	—
Mu Dan Pi	12.54	—
Yu Zhu	12.85	—
Gua Lou	13.81	—
Lian Qiao	14.06	—
Di Gu Pi	14.12	—
Tai Zi Shen	14.82	—
Bai Shao Yao	14.93	—
Ye Jiao Teng	15.09	—
Du Huo	15.10	—
Pi Pa Ye	15.19	—
Chen Pi	15.20	—
Ba Ji Tian	15.24	—
Chai Hu	15.26	—
Jie Geng	16.15	—
Zhe Bei Mu	16.47	—
She Gan	16.63	—
He Huan Pi	17.65	—
Jiao Gu Lan	18.12	—
Jin Yin Hua	18.28	—
Ren Dong Teng	18.37	—
Long Yan Rou	18.64	—
Lu Gen	18.65	—
Shan Yao	18.81	—
Bai Mao Gen	18.84	—
Zhi Ke	18.98	—
(Huai) Niu Xi	20.05	—
Ci Wu Jia	20.23	—
Tian Hua Fen	21.01	—
Gan Jiang	21.05	—
Gui Zhi	21.38	—
Hou Po	21.67	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fo Shou	21.86	—
Tao Ren	21.93	—
Ma Huang	22.00	—
Da Zao	22.07	—
Huang Qi	22.12	—
Chuang Mu Xiang	22.24	—
Chuan Lian Zi	22.41	—
Bai Bian Dou	22.52	—
Ji Li	22.67	—
Shen Qu	23.79	—
Hong Jing Tian	24.28	—
Lian Zi	24.58	—
Bai Zhu	24.69	—
Fang Feng	24.69	—
Huang Qin	24.80	—
Gan Cao	24.81	—
Bai Zi Ren	25.03	—
Qin Jiao	25.19	—
Gou Teng	25.24	—
Sha Shen (Bei)	25.52	—
Huo Ma Ren	25.64	—
Wu Wei Zi	25.67	—
Ji Xue Teng	25.69	—
E Zhu	25.77	—
Bai He	26.01	—
Zi Wan	26.02	—
Rou Gui	26.02	—
Ban Xia (Jiang)	26.23	—
Mu Gua	26.63	—
Mang Xiao	27.15	—
Fu Zi	27.20	—
Sheng Jiang	28.39	—
Tian Men Dong	28.56	—
Tu Fu Ling	28.66	—
Fu Ling	28.68	—
Yin Yang Huo	29.20	—
(Shi) Chang Pu	29.77	—
Fang Ji	29.98	—
Dan Shen	30.07	—
Yan Hu Suo	30.53	—
Lai Fu Zi	31.01	—
Fu Xiao Mai	31.60	—
Suan Zao Ren	31.62	—
Mai Men Dong	31.99	—
Qing Pi	32.24	—
Mao Dong Qing	32.27	—
Mai Ya	32.41	—
Zhu Ru	32.58	—
Bai Xian Pi	33.36	—
Xian He Cao	33.56	—
Ling Zhi	34.06	—
Che Qian Zi	34.29	—
Xie Bai	35.04	—
Wu Yao	35.14	—
Yi Yi Ren	35.83	—
Zi Hua Di Ding	37.13	—
Jiang Huang	37.37	—
Ze Xie	37.72	—
Zi Su Zi	38.14	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ju Hua	38.36	—
Yi Mu Cao	38.89	—
Dan Dou Chi	38.94	—
Xiang Fu	39.81	—
Chi Shao (Yao)	40.30	—
Ban Zhi Lian	40.66	—
Fu Shen	41.91	—
Sha Ren	42.37	—
Qiang Huo	43.12	—
Zhu Ling	43.75	—
Hong Hua	44.00	—
Yu Jin	44.01	—
Huang Lian	44.85	—
Cang Er Zi	46.72	—
Ku Shen	46.85	—
Ma Huang Gen	47.27	—
Niu Bang Zi	47.39	—
(Bai) Dou Kou	48.60	—

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	5.78
62650	62650	0.00	7.54
62414	62414	0.00	5.06
62649	62649	0.00	7.32

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 2.1-2020-01

Validated substance/substance group **Dang Gui Wei**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60308-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dang Gui Wei	2	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	G312H1307621	61885	40	from supplier
PhytoComm	Dang Gui Wei	G312H1307621	61886	40	from supplier
PhytoComm	Dang Gui Wei	G312HS291SK1	62621	40	from supplier
PhytoComm	Dang Gui Wei	G312HS291SK1	62622	40	from supplier

Validation samples

A total of 39 821 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 *Dang Gui Wei*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G312H1307621	61885 [†]	20
PhytoComm	Dang Gui Wei	G312H1307621	61886 [†]	20
PhytoComm	Dang Gui Wei	G312HS291SK1	62621 [†]	20
PhytoComm	Dang Gui Wei	G312HS291SK1	62622 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	160	0	25 641
Type B	1	80	0	13 086
Type C	9	0	0	844

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 % (> 96.2500 %)
Type B	99.9927 % (> 99.9638 %)	100.0000 % (> 92.5000 %)
Type C	99.5384 % (> 98.9546 %)	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 *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.

Substance	Distance in main model	Distance in second-stage model
Bai Mao Gen	6.51	—
Chai Hu	7.19	—
Gou Qi Zi	7.61	—
Chen Pi	7.71	—
Bai Shao Yao	7.92	—
Dang Shen	7.96	—
Yuan Zhi	8.00	—
She Gan	8.20	—
Gua Lou	8.92	—
Yu Zhu	9.43	—
Chuan Xiong	9.92	—
Pi Pa Ye	9.95	—
Cang Zhu	10.27	—
Fo Shou	10.51	—
Dang Gui	10.55	—
San Qi	11.27	—
Zhi Mu	11.36	—
Tai Zi Shen	12.66	—
Ye Jiao Teng	12.83	—
Zhe Bei Mu	13.05	—
Du Huo	13.30	—
Sang Zhi	13.31	—
Shen Qu	13.34	—
Di Gu Pi	14.21	—
Lu Gen	14.32	—
Huang Qi	14.61	—
(Huai) Niu Xi	14.70	—
San Leng	15.11	—
Ren Dong Teng	15.43	—
Jie Geng	15.59	—
Lian Qiao	15.83	—
Da Zao	15.90	—
Shan Yao	15.92	—
He Huan Pi	15.97	—
Zi Wan	16.24	—
Tian Men Dong	16.38	—
Ci Wu Jia	17.25	—
Long Yan Rou	17.32	—
Bai Zi Ren	17.70	—
Tao Ren	17.93	—
Ji Li	17.94	—
Rou Gui	18.01	—
Lian Zi	18.43	—
Bai Bian Dou	18.47	—
Long Dan (Cao)	18.48	—
Bai Zhu	19.17	—
Mu Dan Pi	19.66	—
Shan Yu Rou	19.67	—
Tu Fu Ling	19.81	—
Tian Hua Fen	19.84	—
Gui Zhi	19.86	—
Huo Ma Ren	20.02	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Jiang	21.32	—
Gou Teng	21.34	—
Ba Ji Tian	21.34	—
Fang Feng	21.45	—
Jiao Gu Lan	21.60	—
Fu Ling	22.55	—
Sheng Jiang	22.76	—
Ma Huang	22.83	—
Qing Pi	23.06	—
Ban Xia (Jiang)	23.14	—
Chuang Mu Xiang	23.20	—
Hong Jing Tian	23.24	—
Dan Shen	23.46	—
Zhi Ke	23.59	—
Jin Yin Hua	23.60	—
Mu Gua	23.78	—
Ji Xue Teng	23.99	—
Fu Zi	24.20	—
(Shi) Chang Pu	24.78	—
Fu Xiao Mai	25.06	—
Hou Po	25.40	—
Ling Zhi	25.42	—
Mao Dong Qing	25.75	—
Bai Xian Pi	25.94	—
Xiang Fu	25.97	—
Zhu Ru	26.49	—
Wu Wei Zi	27.20	—
Ju Hua	27.38	—
Huang Qin	27.65	—
Yin Yang Huo	27.68	—
Gan Cao	29.08	—
E Zhu	29.26	—
Mang Xiao	29.29	—
Xie Bai	29.38	—
Chuan Lian Zi	29.77	—
Yi Mu Cao	30.01	—
Suan Zao Ren	30.08	—
Yi Yi Ren	30.45	—
Lai Fu Zi	30.91	—
Wu Yao	31.13	—
Yan Hu Suo	31.31	—
Bai He	31.36	—
Chi Shao (Yao)	31.70	—
Che Qian Zi	31.88	—
Ze Xie	32.64	—
Fang Ji	32.68	—
Qin Jiao	33.55	—
Dan Dou Chi	33.68	—
Ma Huang Gen	35.29	—
Zi Su Zi	35.50	—
Jiang Huang	36.15	—
Sha Shen (Bei)	36.83	—
Ku Shen	37.10	—
Fu Shen	37.41	—
Mai Men Dong	37.44	—
Yu Jin	37.79	—
Niu Bang Zi	38.64	—
Xian He Cao	39.25	—
Qiang Huo	41.88	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Hua Di Ding	42.52	—
Wang Bu Liu Xing	42.88	—
Huang Lian	44.50	—
Sha Ren	44.67	—
Wei Ling Xian	44.73	—
Cang Er Zi	44.77	—
Ban Zhi Lian	46.20	—
(Bai) Dou Kou	46.51	—
Mai Ya	47.11	—
Gu Sui Bu	49.08	—

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
61886	61886	0.00	6.51
61885	61885	0.00	7.19
62622	62622	0.00	7.92
62621	62621	0.00	7.71

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 2.1-2020-01

Validated substance/substance group **Dang Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60000-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Dang Shen; Codonopsis pilosulae radix

Special notes

When selecting the *Dang Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dang Shen	3	0	4

Second-stage model

For differentiation of the substance/substance group *Dang 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 *Dang Shen*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dang Shen	G077H2001821	62367	40	from supplier
PhytoComm	Dang Shen	G077H2001821	62368	40	from supplier
PhytoComm	Dang Shen	G077HS352RQ1	62369	40	from supplier
PhytoComm	Dang Shen	G077HS352RQ1	62370	40	from supplier
PhytoComm	Dang Shen	G077HS352SG1	62673	40	from supplier
PhytoComm	Dang Shen	G077HS352SG1	62674	40	from supplier

Validation samples

A total of 39 821 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 *Dang Shen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *Dang 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	Dang Shen	G077H2001821	62367 [†]	20
PhytoComm	Dang Shen	G077H2001821	62368 [†]	20
PhytoComm	Dang Shen	G077HS352RQ1	62369 [†]	20
PhytoComm	Dang Shen	G077HS352RQ1	62370 [†]	20
PhytoComm	Dang Shen	G077HS352SG1	62673 [†]	20
PhytoComm	Dang Shen	G077HS352SG1	62674 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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 4 batches from the substance/substance group *Dang 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
EuRho	Dang Shen	g077h2001222	2
Phytocomm	Dang Shen	G077H2001022	1
Phytocomm	Dang Shen	G077H2001521	2
PhytoComm	Dang Shen	G077H2001521	2
PhytoComm	Dang Shen	G077H2001621	1

- 845 spectra from 12 *Apo-Ident* customers from a total of 512 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 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 *Dang 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	240	0	25 561
Type B	1	117	3	13 046
Type C	1	4	4	844

The substance/substance group *Dang 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 % (> 97.5000 %)
Type B	99.9855 % (> 99.9565 %)	97.5000 % (> 95.0000 %)
Type C	99.9612 % (> 99.3727 %)	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 *Dang 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
Yuan Zhi	5.45	—
Zhi Mu	8.10	—
Dang Gui Wei	8.41	—
Bai Shao Yao	8.45	—
Chen Pi	8.48	—
Fo Shou	8.83	—
Tian Men Dong	9.75	—
Chuan Xiong	10.16	—
Jie Geng	11.00	—
Pi Pa Ye	11.26	—
Dang Gui	12.97	—
Di Gu Pi	12.97	—
Gua Lou	13.12	—
San Qi	13.14	—
Huang Qi	13.20	—
Ye Jiao Teng	13.26	—
Qin Jiao	13.61	—
Shen Qu	13.87	—
Bai Zhu	14.82	—
Yu Zhu	15.20	—
Cang Zhu	15.32	—
Mu Gua	15.38	—
Shan Yao	15.69	—
Rou Gui	16.23	—
Da Zao	16.55	—
He Huan Pi	16.78	—
Gou Qi Zi	16.95	—
Tai Zi Shen	17.43	—
San Leng	17.61	—
Shan Yu Rou	17.68	—
She Gan	17.99	—
Ji Li	18.17	—
Lian Zi	18.22	—
Ci Wu Jia	18.75	—
Bai Bian Dou	18.83	—
Gui Zhi	19.18	—
Ren Dong Teng	19.52	—
Chuang Mu Xiang	19.84	—
Chai Hu	20.02	—
Bai Zi Ren	20.26	—
Tu Fu Ling	20.30	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Dan (Cao)	20.31	—
Zhe Bei Mu	20.45	—
Lian Qiao	20.50	—
Tao Ren	20.85	—
Fu Ling	21.35	—
Ban Xia (Jiang)	21.40	—
Mu Dan Pi	21.75	—
Bai He	22.27	—
Tian Hua Fen	22.57	—
Long Yan Rou	22.88	—
Fu Xiao Mai	22.96	—
Huo Ma Ren	23.13	—
Sheng Jiang	23.93	—
Sang Zhi	24.18	—
Gou Teng	25.04	—
Zhu Ru	25.29	—
(Huai) Niu Xi	25.38	—
Mai Men Dong	25.69	—
Jiao Gu Lan	26.31	—
Ling Zhi	26.36	—
Chi Shao (Yao)	26.36	—
Bai Xian Pi	27.05	—
Mao Dong Qing	27.26	—
Ji Xue Teng	28.06	—
Dan Shen	28.23	—
Bai Mao Gen	28.66	—
Suan Zao Ren	28.71	—
Zhi Ke	28.87	—
Mang Xiao	29.99	—
Jin Yin Hua	30.33	—
Ma Huang	30.69	—
Gan Cao	30.86	—
Huang Qin	30.97	—
Fu Zi	31.12	—
Yi Yi Ren	31.29	—
Zi Wan	31.82	—
Wu Wei Zi	31.85	—
Hou Po	32.01	—
Hong Jing Tian	32.20	—
Du Huo	32.36	—
Ba Ji Tian	33.22	—
Ze Xie	33.27	—
Lai Fu Zi	33.99	—
Gan Jiang	34.57	—
Yin Yang Huo	34.57	—
Wang Bu Liu Xing	34.99	—
Che Qian Zi	36.55	—
Yan Hu Suo	36.87	—
Qing Pi	36.99	—
Fang Feng	37.25	—
Fang Ji	37.29	—
Chuan Lian Zi	37.62	—
Ma Huang Gen	38.17	—
Fu Shen	38.74	—
Dan Dou Chi	39.81	—
(Shi) Chang Pu	40.92	—
Yu Jin	41.08	—
Lu Gen	41.82	—
E Zhu	44.90	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Shen (Bei)	47.80	—
Ban Zhi Lian	47.90	—
Mai Ya	48.33	—
Xiang Fu	49.25	—
Cang Er Zi	49.46	—
Qiang Huo	49.89	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dang 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
62369	62369	0.00	8.10
62674	62674	0.00	5.45
62368	62368	0.00	11.00
62673	62673	0.00	5.68
62367	62367	0.00	11.02
62370	62370	0.00	8.26

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 2.1-2020-01

Validated substance/substance group **Di Fu Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002264-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	1	852

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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
Guang Huo Xiang	30.21	—
Hua Shi	45.59	—
Xin Yi	50.09	—

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
62738	62738	0.00	30.59
62737	62737	0.00	30.21

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 2.1-2020-01

Validated substance/substance group **Di Gu Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60036-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G149H0601622	62211	40	from supplier
PhytoComm	Di Gu Pi	G149H0601622	62212	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G149H0601622	62211 [†]	20
PhytoComm	Di Gu Pi	G149H0601622	62212 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	G149H0601421	1
Phytocomm	Di Gu Pi	g149h060112	1
PhytoComm	Di Gu Pi	G149H0601421	2

- 849 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 *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	1	237	3	25 560
Type B	2	112	8	13 045
Type C	5	1	3	844

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	99.9964 % (> 99.9814 %)	98.7500 % (> 97.5000 %)
Type B	99.9855 % (> 99.9565 %)	93.3333 % (> 90.8333 %)
Type C	99.6331 % (> 99.0454 %)	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
Ban Xia (Jiang)	3.71	—
He Huan Pi	4.80	—
Sheng Jiang	5.19	—
Gui Zhi	5.19	—
Ci Wu Jia	6.54	—
Lian Zi	7.19	—
Rou Gui	7.54	—
Zhu Ru	8.97	—
Bai Shao Yao	9.08	—
Sang Zhi	9.14	—
Tao Ren	9.43	—
Zhe Bei Mu	10.30	—
Bai Bian Dou	10.61	—
Jie Geng	10.84	—
Shen Qu	10.92	—
Bai Zi Ren	11.01	—
Yi Yi Ren	11.46	—
Bai Zhu	11.71	—
Fu Xiao Mai	11.91	—
San Leng	12.10	—
Ji Li	12.42	—
Tai Zi Shen	12.76	—
Gou Teng	13.63	—
Fu Ling	13.81	—
Huo Ma Ren	14.83	—
Tian Hua Fen	15.44	—
Ling Zhi	15.63	—
Fo Shou	16.04	—
Ye Jiao Teng	16.35	—
Bai Xian Pi	16.61	—
Gua Lou	16.66	—
Ji Xue Teng	16.74	—
Qing Pi	17.27	—
Dang Shen	17.42	—
Ze Xie	17.78	—
She Gan	18.00	—
Gou Qi Zi	18.32	—
Dang Gui Wei	18.56	—
Huang Qi	18.76	—
Pi Pa Ye	18.80	—
Yuan Zhi	19.44	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Dan (Cao)	19.58	—
Chuan Xiong	19.93	—
Qin Jiao	20.15	—
Tian Men Dong	20.35	—
Ma Huang Gen	20.81	—
Long Yan Rou	20.98	—
Du Huo	21.11	—
Chen Pi	21.42	—
San Qi	21.74	—
Mu Dan Pi	21.84	—
Shan Yao	22.19	—
Lian Qiao	22.20	—
Bai Mao Gen	22.25	—
Mu Gua	22.70	—
Da Zao	23.05	—
Gan Jiang	23.11	—
Tu Fu Ling	23.46	—
Fu Shen	23.72	—
Wang Bu Liu Xing	23.86	—
(Shi) Chang Pu	24.04	—
Chuan Lian Zi	24.43	—
Chi Shao (Yao)	24.62	—
Cang Zhu	24.85	—
Ren Dong Teng	24.92	—
Lai Fu Zi	25.05	—
Bai He	25.06	—
Dang Gui	25.42	—
Yan Hu Suo	25.56	—
Lu Gen	25.63	—
Yu Jin	25.79	—
Chai Hu	26.46	—
Zhi Mu	26.80	—
Cang Er Zi	27.02	—
Mai Men Dong	27.74	—
Fu Zi	28.57	—
Qiang Huo	29.07	—
Wu Wei Zi	29.24	—
Mang Xiao	29.26	—
Zi Wan	29.56	—
Dan Shen	29.59	—
Jin Yin Hua	29.80	—
Shan Yu Rou	29.83	—
Zhi Ke	31.08	—
Xiang Fu	31.67	—
(Huai) Niu Xi	31.74	—
Niu Bang Zi	31.94	—
Jiao Gu Lan	32.07	—
Dan Dou Chi	32.08	—
Zi Su Zi	32.54	—
Mao Dong Qing	32.62	—
Ban Zhi Lian	32.98	—
Ku Shen	33.20	—
Yi Mu Cao	33.50	—
Fang Ji	33.56	—
E Zhu	33.58	—
Yu Zhu	33.66	—
Huang Qin	34.63	—
Yin Yang Huo	35.00	—
Ma Huang	35.19	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ba Ji Tian	35.33	—
Gan Cao	36.06	—
Hou Po	36.80	—
Mai Ya	37.27	—
Huang Lian	38.11	—
Jiang Huang	38.67	—
Che Qian Zi	39.26	—
Sha Shen (Bei)	40.97	—
Hong Jing Tian	42.96	—
Wu Yao	43.55	—
Pu Gong Ying	43.68	—
Suan Zao Ren	44.25	—
Dong Gua Zi	45.04	—
Fang Feng	45.06	—
Zi Hua Di Ding	45.29	—
Xie Bai	45.39	—
Du Zhong	45.88	—
Zhi Zi	47.32	—
Wei Ling Xian	48.59	—
(Bai) Dou Kou	49.13	—
Zhu Ling	49.82	—

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
62692	62692	0.00	3.71
62348	62348	0.00	9.44
62212	62212	0.00	10.51
62347	62347	0.00	9.14
62211	62211	0.00	10.84
62691	62691	0.00	3.94

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 2.1-2020-01

Validated substance/substance group **Dong Gua Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60202-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	1	852

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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
Suan Zao Ren	18.06	—
Cang Er Zi	20.00	—
Ku Shen	27.61	—
Zhi Zi	28.37	—
Shan Yao	29.05	—
Shen Qu	32.58	—
Huo Ma Ren	33.76	—
Wei Ling Xian	35.74	—
Mang Xiao	38.44	—
San Qi	39.80	—
Gan Jiang	41.09	—
Sang Zhi	42.45	—
Di Gu Pi	43.26	—
Gua Lou	43.62	—
Bai Bian Dou	44.15	—
Mu Gua	44.26	—
Wu Wei Zi	45.74	—
Zi Su Zi	47.22	—
She Gan	48.26	—
Gan Cao	48.63	—
Bai Shao Yao	48.89	—
Jie Geng	49.75	—

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
62514	62514	0.00	18.06
62513	62513	0.00	18.37

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 2.1-2020-01

Validated substance/substance group **Dong Kui Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60483-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Dong Kui Zi; Abutili semen

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Dong Kui Zi	1	0	0

Second-stage model

For differentiation of the substance/substance group *Dong Kui 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 Kui Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Dong Kui Zi	G001H0834821	62505	40	from supplier
PhytoComm	Dong Kui Zi	G001H0834821	62506	40	from supplier

Validation samples

A total of 39 821 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 Kui Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Kui 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 Kui Zi	G001H0834821	62505 [†]	20
PhytoComm	Dong Kui Zi	G001H0834821	62506 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Dong Kui 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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 *Dong Kui 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 Kui 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	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

The substance/substance group *Dong Kui 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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 *Dong Kui 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
Shen Qu	23.38	—
Cang Er Zi	26.87	—
Gui Zhi	46.33	—
Shan Yao	46.89	—
Zhe Bei Mu	49.19	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Dong Kui 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
62506	62506	0.00	23.38
62505	62505	0.00	25.44

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 2.1-2020-01

Validated substance/substance group **Du Huo**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50259-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G021H1610821	62391	40	from supplier
PhytoComm	Du Huo	G021H1610821	62392	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G021H1610821	62391 [†]	20
PhytoComm	Du Huo	G021H1610821	62392 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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	g021h610221	1
Phytocomm	Du Huo	G021H1610421	2

- 850 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 *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	25 721
Type B	0	40	0	13 127
Type C	1	0	3	849

The substance/substance group *Du 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.9225 % (> 99.3353 %)	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
Lu Gen	9.77	—
Mu Dan Pi	10.17	—
Gou Qi Zi	10.33	—
Dang Gui Wei	11.99	—
Dang Gui	12.00	—
She Gan	12.74	—
Chai Hu	12.75	—
Bai Mao Gen	14.33	—
Chuan Xiong	16.03	—
(Shi) Chang Pu	17.00	—
Zi Wan	17.07	—
Fang Feng	17.10	—
Gui Zhi	17.37	—
San Leng	17.65	—
Zhi Mu	19.33	—
Cang Zhu	19.44	—
Long Dan (Cao)	20.37	—
Huang Qi	20.91	—
Hou Po	21.11	—
Ba Ji Tian	21.19	—
Sang Zhi	21.38	—
Hong Jing Tian	22.54	—
Ju Hua	22.94	—
Wu Yao	23.25	—
Gua Lou	23.26	—
Gan Jiang	23.27	—
Lian Qiao	23.30	—
Chen Pi	23.63	—
Qing Pi	24.10	—
Bai Shao Yao	24.57	—
(Huai) Niu Xi	24.64	—
Di Gu Pi	24.90	—
Ji Li	25.76	—
Yi Mu Cao	27.18	—
Xiang Fu	27.54	—
Xian He Cao	28.12	—
Jie Geng	28.13	—
Dang Shen	29.11	—
Mang Xiao	29.38	—
E Zhu	31.36	—
Tu Fu Ling	31.75	—
Xie Bai	31.89	—
Tian Hua Fen	32.12	—
Bai Zhu	32.75	—
Jiang Huang	33.25	—
Mu Gua	34.07	—
He Huan Pi	34.43	—
Qin Jiao	34.81	—
Zhi Ke	35.92	—
Chuan Lian Zi	36.42	—
Sha Ren	36.44	—
Huo Ma Ren	36.55	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Shen (Bei)	36.99	—
Yu Jin	37.94	—
Lai Fu Zi	38.23	—
Chi Shao (Yao)	39.39	—
Zi Su Zi	39.40	—
Ku Shen	40.15	—
Da Zao	40.27	—
Gu Sui Bu	40.51	—
Yan Hu Suo	40.56	—
Shan Yao	40.82	—
Huang Jing	41.94	—
Sheng Jiang	42.18	—
Qiang Huo	42.81	—
Wei Ling Xian	42.99	—
Yuan Zhi	43.11	—
He Shou Wu	43.33	—
Ze Xie	43.45	—
(Bai) Dou Kou	43.63	—
Ge Gen	44.36	—
Bai He	44.80	—
Niu Bang Zi	45.47	—
Ma Bian Cao	45.55	—
Wu Wei Zi	46.10	—
Huang Bo	47.69	—
Sang Ji Shend	47.84	—
Xu Duan	48.17	—
Mai Men Dong	48.43	—
Cang Er Zi	48.51	—
Sang Bai Pi	50.21	—

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
62391	62391	0.00	9.77
62392	62392	0.00	9.80

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 2.1-2020-01

Validated substance/substance group **Du Zhong**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60214-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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

- 13 047 spectra from a total of 317 batches from further 172 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	g101h076521	1
Phytocomm	Du Zhong	G101H0736522	5

- 845 spectra from 12 *Apo-Ident* customers from a total of 512 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	236	4	25 561
Type B	0	109	11	13 047
Type C	0	0	8	845

The substance/substance group *Du Zhong* 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 %)	98.3333 % (> 97.0833 %)
Type B	100.0000 % (> 99.9421 %)	90.8333 % (> 88.3333 %)
Type C	100.0000 % (> 98.8230 %)	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
Yin Chen Hao	3.84	—
Yu Jin	4.35	—
Mao Dong Qing	4.73	—
Fang Ji	8.41	—
Ling Zhi	8.57	—
Qiang Huo	8.66	—
Hou Po	10.55	—
Bai Xian Pi	10.79	—
Yin Yang Huo	11.46	—
Zi Hua Di Ding	11.72	—
Bu Gu Zhi	11.99	—
Yan Hu Suo	13.07	—
Sang Bai Pi	13.09	—
Jing Jie	13.57	—
Dan Zhu Ye	13.67	—
Jiao Gu Lan	14.22	—
Dan Dou Chi	14.44	—
Qing Pi	14.70	—
Shen Qu	14.89	—
Shan Yao	15.10	—
Mu Dan Pi	15.33	—
Ma Bian Cao	15.57	—
Bai Bian Dou	15.63	—
Che Qian Zi	16.55	—
Chai Hu	17.46	—
Ye Jiao Teng	17.90	—
Xu Duan	17.97	—
Huang Bai	18.17	—
Ji Li	18.19	—
Pi Pa Ye	18.43	—
Tu Fu Ling	18.80	—
Ren Dong Teng	19.06	—
Chuan Xiong	19.32	—
Wu Jia Pi	19.54	—
Dan Shen	19.87	—
Lian Zi	20.05	—
Pu Gong Ying	20.18	—
Sha Ren	20.84	—
Hong Jing Tian	21.25	—
Yuan Zhi	21.46	—
Huang Lian	21.57	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xi Xian Cao	21.60	—
Nü Zhen Zi	22.24	—
Ji Xue Teng	22.50	—
Jin Yin Hua	24.29	—
Ma Huang Gen	24.80	—
Ma Huang	24.93	—
Zhi Ke	25.08	—
Lian Qiao	25.41	—
Ban Xia (Jiang)	26.13	—
He Shou Wu	26.30	—
Sang Ji Shend	26.40	—
(Sheng) Di Huang	26.53	—
Bai Shao Yao	26.77	—
He Huan Pi	27.36	—
Wu Wei Zi	27.61	—
Rou Gui	27.69	—
Gua Lou	27.78	—
Gou Teng	28.20	—
Di Gu Pi	29.48	—
Xian He Cao	29.96	—
Sheng Jiang	30.40	—
Cang Zhu	30.45	—
Jin Qian Cao	30.97	—
Fu Xiao Mai	31.25	—
Chen Pi	31.25	—
Sang Ye	31.37	—
Wang Bu Liu Xing	31.38	—
Fu Zi	31.38	—
Gui Zhi	32.82	—
Xia Ku Cao	32.88	—
Gou Qi Zi	33.11	—
Yi Mu Cao	33.16	—
Yi Yi Ren	33.55	—
Xuan Shen	33.75	—
Wu Zhu Yu	35.16	—
Mang Xiao	35.30	—
Huo Ma Ren	35.65	—
Da Zao	35.69	—
Wu Mei	35.84	—
Cang Er Zi	36.44	—
Gu Sui Bu	37.02	—
Tian Hua Fen	37.24	—
Jiang Huang	37.44	—
E Zhu	37.48	—
Tao Ren	37.72	—
Fu Ling	37.88	—
Zhi Shi	39.31	—
Ban Zhi Lian	39.44	—
Qing Hao	40.11	—
Shan Yu Rou	40.11	—
Huang Bo	40.59	—
Zhe Bei Mu	40.88	—
Xin Yi	41.30	—
Ze Xie	41.77	—
Ge Gen	41.85	—
Sang Zhi	41.91	—
Hong Hua	42.12	—
Jie Geng	42.56	—
Yu Xing Cao	42.71	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Zi Ren	43.14	—
Tai Zi Shen	43.63	—
Fu Shen	43.89	—
Chi Shao (Yao)	44.22	—
San Qi	44.99	—
Suan Zao Ren	45.01	—
Ci Wu Jia	45.63	—
Zhu Ru	46.59	—
Bai He	47.63	—
Huang Qin	47.83	—
Bai Zhu	48.13	—
Zi Su Zi	48.23	—
Sha Shen (Bei)	48.61	—
Long Yan Rou	48.82	—
Bai Mao Gen	48.84	—
Hua Shi	48.93	—
Chuang Mu Xiang	49.12	—
Zhi Mu	49.22	—
Hu Zhang	49.46	—
Fang Feng	49.78	—
Han Lian Cao	51.38	—

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
62580	62580	0.00	4.75
62729	62729	0.00	4.35
62579	62579	0.00	4.73
62336	62336	0.00	3.84
62335	62335	0.00	4.15
62730	62730	0.00	4.46

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 2.1-2020-01

Validated substance/substance group **E Zhu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50344-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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	h1110021	1
Phytocomm	E Zhu	H1110021	1
PhytoComm	E Zhu	G083H1110222	2

- 849 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	1	0	4	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
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.

[†]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	15.57	—
Cang Er Zi	15.89	—
Jiang Huang	17.72	—
Sha Shen (Bei)	19.05	—
Gui Zhi	21.57	—
Tu Fu Ling	22.30	—
Sha Ren	22.46	—
Mang Xiao	22.61	—
Sang Zhi	22.84	—
(Bai) Dou Kou	23.12	—
He Huan Pi	23.66	—
Yan Hu Suo	24.37	—
Chuan Xiong	24.53	—
Chuan Lian Zi	25.87	—
Bai Bian Dou	26.71	—
Mo Yao	27.33	—
Niu Bang Zi	29.43	—
Wang Bu Liu Xing	29.46	—
Zi Su Zi	29.64	—
Long Dan (Cao)	29.66	—
Ba Ji Tian	30.00	—
Bai He	30.08	—
Lai Fu Zi	30.80	—
Chai Hu	30.96	—
Chen Pi	31.18	—
Mu Dan Pi	31.32	—
Sheng Jiang	31.43	—
San Qi	31.95	—
She Gan	31.99	—
Zhu Ling	32.91	—
Shan Yao	33.16	—
Di Gu Pi	33.23	—
Mai Ya	34.86	—
Du Huo	35.49	—
Yi Mu Cao	36.53	—
Xie Bai	37.43	—
Yu Jin	38.16	—
Gan Jiang	38.30	—
Huo Ma Ren	38.51	—
Xiang Fu	38.93	—
(Huai) Niu Xi	38.96	—
Yuan Zhi	39.46	—
Ban Xia (Jiang)	39.63	—
Gou Qi Zi	39.75	—
Wu Yao	40.04	—
Hou Po	40.49	—
Fu Ling	40.59	—
Huang Lian	41.06	—
Dang Gui	41.13	—
Xian He Cao	41.13	—
Lu Gen	41.61	—
Cang Zhu	42.40	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fang Feng	43.31	—
Rou Gui	43.55	—
Jie Geng	43.76	—
Tian Hua Fen	44.28	—
Dang Shen	44.63	—
Da Zao	45.13	—
Fang Ji	45.36	—
(Shi) Chang Pu	45.69	—
Qing Pi	45.79	—
Zhi Mu	45.97	—
Ban Zhi Lian	46.38	—
Xin Yi	46.66	—
San Leng	47.08	—
Zhi Ke	47.45	—
Qiang Huo	47.69	—
Gua Lou	47.91	—
Hong Jing Tian	48.05	—
Lian Zi	48.85	—
Bai Zhu	49.50	—
Yin Yang Huo	49.71	—
Dan Dou Chi	49.72	—
Qin Jiao	49.95	—

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	15.57
62516	62516	0.00	15.89

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 2.1-2020-01

Validated substance/substance group **Fang Feng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50261-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G217H0728821	62307	40	from supplier
PhytoComm	Fang Feng	G217H0728821	62308	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G217H0728821	62307 [†]	20
PhytoComm	Fang Feng	G217H0728821	62308 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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	g217h0728923	1
Phytocomm	Fang Feng	g217h0728221	1
Phytocomm	Fang Feng	G217H0728521	3
PhytoComm	Fang Feng	G217H0728521	2

- 846 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	2	0	7	844

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.7674 % (> 99.1791 %)	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	9.70	—
Gou Qi Zi	10.56	—
Gui Zhi	12.27	—
Ju Hua	13.24	—
Du Huo	14.48	—
Chai Hu	17.87	—
Dang Gui Wei	18.16	—
Lu Gen	18.20	—
Chuan Xiong	18.62	—
Zi Wan	18.90	—
Hou Po	19.08	—
Huang Jing	19.58	—
(Shi) Chang Pu	19.78	—
Wu Yao	20.32	—
Dang Gui	20.97	—
She Gan	21.02	—
Ba Ji Tian	22.21	—
Zhi Mu	22.49	—
Gan Jiang	22.66	—
Yi Mu Cao	23.07	—
Lian Qiao	23.08	—
Mu Dan Pi	23.47	—
Hong Jing Tian	23.49	—
Cang Zhu	23.89	—
Xie Bai	24.00	—
He Huan Pi	24.12	—
Gua Lou	24.13	—
Bai Mao Gen	24.97	—
Ji Li	26.09	—
Gu Sui Bu	26.11	—
Huang Qi	26.91	—
San Leng	27.13	—
Qiang Huo	27.62	—
Mang Xiao	27.79	—
Sang Zhi	27.97	—
Xu Duan	28.32	—
Qing Pi	29.22	—
Jiang Huang	29.38	—
Xian He Cao	29.79	—
Long Dan (Cao)	31.85	—
Huo Ma Ren	32.27	—
Ma Bian Cao	32.30	—
Yu Jin	34.11	—
(Huai) Niu Xi	34.30	—
Dan Shen	34.39	—
Jie Geng	34.69	—
Sang Ji Shend	34.70	—
E Zhu	34.71	—
Sha Ren	34.93	—
Di Gu Pi	34.95	—
Xiang Fu	35.08	—
Dang Shen	35.60	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
He Shou Wu	35.67	—
Tu Fu Ling	36.68	—
Huang Bo	36.80	—
Sang Bai Pi	37.54	—
Ku Shen	37.63	—
Wu Zhu Yu	37.79	—
Zi Su Zi	38.31	—
Mu Gua	39.40	—
Yin Chen Hao	39.82	—
Zhi Ke	40.47	—
Lai Fu Zi	40.53	—
Yan Hu Suo	41.04	—
Chuan Lian Zi	41.33	—
Wei Ling Xian	41.50	—
Nü Zhen Zi	41.74	—
Sheng Jiang	41.88	—
Yin Yang Huo	42.42	—
Tian Hua Fen	42.51	—
Da Zao	43.34	—
Shan Yao	43.62	—
Bai Zhu	43.84	—
Niu Bang Zi	44.02	—
Ge Gen	44.86	—
Xin Yi	45.26	—
(Bai) Dou Kou	45.47	—
Fu Zi	45.69	—
Ze Xie	46.95	—
Chi Shao (Yao)	47.24	—
Cang Er Zi	47.30	—
Yuan Zhi	47.44	—
Bai Shao Yao	47.61	—
Bai Bian Dou	48.23	—

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
62308	62308	0.00	10.53
62307	62307	0.00	9.70

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 2.1-2020-01

Validated substance/substance group **Fang Ji**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60270-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Fang Ji; Stephaniae tetrandrae radix

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Fang Ji	1	0	2

Second-stage model

For differentiation of the substance/substance group *Fang Ji* 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 Ji*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Fang Ji	G236HS122QM1	62179	40	from supplier
PhytoComm	Fang Ji	G236HS122QM1	62180	40	from supplier

Validation samples

A total of 39 821 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 Ji*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Ji*.
- 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 Ji	G236HS122QM1	62179 [†]	20
PhytoComm	Fang Ji	G236HS122QM1	62180 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Fang Ji*.
- 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	Fang Ji	G236H0727321	1
PhytoComm	Fang Ji	G236H0727123	1

- 851 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 *Fang Ji* 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 Ji* 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	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

The substance/substance group *Fang Ji* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 Ji* 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 Pi	4.45	—
Jiao Gu Lan	8.54	—
Huang Bai	10.24	—
Zi Hua Di Ding	10.78	—
Du Zhong	11.20	—
Dan Shen	12.23	—
Chai Hu	12.90	—
Qiang Huo	13.39	—
Mao Dong Qing	13.84	—
Shan Yao	15.88	—
Yan Hu Suo	16.19	—
Wu Wei Zi	16.54	—
Sang Ye	17.19	—
Bai Bian Dou	17.66	—
Pi Pa Ye	18.09	—
Hong Jing Tian	18.98	—
Jin Yin Hua	19.30	—
Mu Dan Pi	19.68	—
Zhi Ke	20.08	—
Ling Zhi	20.43	—
Yu Jin	20.58	—
Yin Yang Huo	20.89	—
Hou Po	20.90	—
Ji Li	20.99	—
Dan Dou Chi	21.02	—
Chuan Xiong	21.72	—
Bai Xian Pi	22.44	—
Huang Lian	22.73	—
Ren Dong Teng	23.33	—
Shan Yu Rou	23.69	—
Gua Lou	24.04	—
Bai Shao Yao	24.30	—
Xi Xian Cao	24.36	—
Lian Qiao	24.71	—
Hong Hua	25.08	—
Cang Zhu	25.51	—
Ma Huang	25.60	—
Ye Jiao Teng	26.22	—
Che Qian Zi	26.54	—
Wu Mei	26.74	—
Ji Xue Teng	27.31	—
Shen Qu	27.50	—
Di Gu Pi	27.75	—
He Huan Pi	27.87	—
Lian Zi	27.98	—
Bai Zhu	28.61	—
Gou Qi Zi	28.68	—
Fu Zi	28.78	—
Cang Er Zi	29.05	—
Gui Zhi	29.16	—
Tu Fu Ling	30.18	—
Pu Gong Ying	30.23	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	31.71	—
Suan Zao Ren	33.57	—
Xia Ku Cao	34.63	—
Ban Xia (Jiang)	34.70	—
Sheng Jiang	36.06	—
Bai Zi Ren	36.66	—
Gou Teng	37.21	—
Jiang Huang	38.11	—
Huo Ma Ren	38.15	—
Chen Pi	38.39	—
Yi Mu Cao	38.70	—
Tao Ren	38.78	—
Dang Gui	38.89	—
San Qi	38.98	—
Dang Shen	39.04	—
Huang Qin	39.77	—
Nü Zhen Zi	40.11	—
Mang Xiao	40.29	—
Zhe Bei Mu	40.34	—
Tai Zi Shen	40.42	—
Fu Xiao Mai	40.55	—
Tian Hua Fen	40.79	—
Chi Shao (Yao)	41.37	—
Bai Mao Gen	41.70	—
Rou Gui	42.93	—
Ze Xie	43.93	—
Fu Ling	44.06	—
Wang Bu Liu Xing	44.88	—
Ma Huang Gen	45.13	—
Yu Zhu	45.16	—
Xin Yi	45.41	—
Sang Zhi	45.97	—
Chuang Mu Xiang	46.14	—
Da Zao	46.20	—
Zhi Mu	47.23	—
Gan Cao	47.40	—
Yi Yi Ren	47.70	—
Jie Geng	48.92	—
Ban Zhi Lian	49.97	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Fang Ji* 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
62180	62180	0.00	4.45
62179	62179	0.00	4.50

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 2.1-2020-01

Validated substance/substance group **Fo Shou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60119-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 720
Type B	0	39	1	13 127
Type C	0	0	0	853

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 % (> 99.9424 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8323 %)	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	4.73	—
Chen Pi	7.35	—
Tian Men Dong	9.26	—
Ye Jiao Teng	9.48	—
Tao Ren	10.79	—
Rou Gui	11.90	—
Bai Shao Yao	12.59	—
Tai Zi Shen	13.25	—
He Huan Pi	13.38	—
Dang Shen	13.41	—
Di Gu Pi	13.58	—
Lian Zi	13.77	—
Yu Zhu	14.34	—
Da Zao	14.54	—
Zhe Bei Mu	14.57	—
Long Yan Rou	14.94	—
Zhi Mu	15.21	—
Shen Qu	15.32	—
Gui Zhi	15.71	—
Ci Wu Jia	16.59	—
Yuan Zhi	16.65	—
Pi Pa Ye	16.66	—
Ban Xia (Jiang)	17.12	—
Sheng Jiang	17.32	—
Dang Gui Wei	17.65	—
Chuan Xiong	17.90	—
Tu Fu Ling	17.91	—
Fu Xiao Mai	17.97	—
Bai Xian Pi	18.64	—
Bai Bian Dou	18.82	—
Gou Teng	18.88	—
Ji Li	19.15	—
Bai Zi Ren	19.31	—
Fu Ling	19.59	—
San Qi	21.17	—
Ji Xue Teng	21.20	—
Ma Huang	21.90	—
Huo Ma Ren	22.17	—
Cang Zhu	23.08	—
Ren Dong Teng	23.22	—
Zhu Ru	23.91	—
Yi Yi Ren	24.08	—
Lian Qiao	24.10	—
Dan Shen	24.98	—
Ling Zhi	25.10	—
Shan Yao	25.16	—
Dang Gui	25.39	—
Shan Yu Rou	26.05	—
Tian Hua Fen	26.20	—
Zhi Ke	26.61	—
Ma Huang Gen	26.77	—
Fu Shen	26.98	—
Gou Qi Zi	28.07	—
Bai Zhu	28.50	—
Fu Zi	29.12	—
Jiao Gu Lan	29.35	—
Jie Geng	29.78	—
Huang Qin	29.95	—
Yin Yang Huo	29.98	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mao Dong Qing	31.47	—
Hong Jing Tian	32.85	—
Yan Hu Suo	33.52	—
Dan Dou Chi	33.62	—
Jin Yin Hua	35.32	—
Ze Xie	35.49	—
Yu Jin	35.83	—
Wu Wei Zi	36.36	—
Sang Zhi	36.82	—
Chai Hu	37.71	—
Hou Po	37.91	—
Chuang Mu Xiang	37.97	—
Fang Ji	38.23	—
Mang Xiao	38.26	—
Gan Cao	38.31	—
Wang Bu Liu Xing	41.60	—
Che Qian Zi	41.83	—
Chuan Lian Zi	41.91	—
Suan Zao Ren	41.97	—
Mai Men Dong	43.46	—
Chi Shao (Yao)	43.69	—
San Leng	44.68	—
Mu Dan Pi	45.81	—
Huang Qi	45.95	—
Huang Lian	46.63	—
Qiang Huo	47.49	—

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
62420	62420	0.00	4.73
62419	62419	0.00	6.82

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 2.1-2020-01

Validated substance/substance group **Fu Ling**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50260-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Fu Ling	2	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

Validation samples

A total of 39 821 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 Ling*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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

- 13 087 spectra from a total of 318 batches from further 172 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	G200H1019321	1
Phytocomm	Fu Ling	g200h1019321	3
Phytocomm	Fu Ling	G200H1019522	2
PhytoComm	Fu Ling	G200H1019221	1
PhytoComm	Fu Ling	G200H1019321	1
PhytoComm	Fu Ling	G200H1019522	1
PhytoComm	Fu Ling	G200H1019622	1

- 841 spectra from 11 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	12	841

The substance/substance group *Fu 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8224 %)	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 *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	8.80	—
Zhu Ru	8.97	—
Gui Zhi	9.37	—
Sheng Jiang	9.44	—
Rou Gui	11.46	—
Ban Xia (Jiang)	11.53	—
Bai Bian Dou	11.92	—
Gou Teng	12.09	—
Lian Zi	12.74	—
Wang Bu Liu Xing	13.16	—
Ji Xue Teng	13.22	—
Ma Huang Gen	14.46	—
Fu Xiao Mai	14.83	—
He Huan Pi	14.97	—
Fu Shen	15.05	—
Ci Wu Jia	15.25	—
Ji Li	15.97	—
Bai Xian Pi	16.00	—
Ren Dong Teng	16.30	—
Ling Zhi	16.59	—
Tao Ren	19.15	—
Huo Ma Ren	19.36	—
Tai Zi Shen	20.62	—
Di Gu Pi	21.68	—
Bai Zi Ren	21.94	—
Bai Shao Yao	22.34	—
Zhe Bei Mu	23.42	—
Ye Jiao Teng	24.46	—
Shen Qu	24.91	—
Dan Dou Chi	25.20	—
San Qi	25.46	—
Tian Men Dong	25.95	—
Yu Jin	26.02	—
Fo Shou	26.03	—
Mao Dong Qing	26.62	—
Chen Pi	26.77	—
Tian Hua Fen	29.21	—
Dan Shen	29.67	—
Long Yan Rou	29.96	—
Gua Lou	30.41	—
Yan Hu Suo	31.35	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yao	31.37	—
Da Zao	31.57	—
Tu Fu Ling	31.88	—
Yuan Zhi	31.89	—
Chuan Xiong	32.39	—
Lian Qiao	32.62	—
Chai Hu	32.67	—
Yin Yang Huo	33.08	—
Zhi Mu	33.15	—
Mang Xiao	33.99	—
Pi Pa Ye	35.96	—
Jiao Gu Lan	37.24	—
Fang Ji	38.36	—
Dang Gui Wei	39.89	—
Dang Shen	40.83	—
Che Qian Zi	40.96	—
Hong Jing Tian	41.05	—
Fu Zi	41.29	—
Bai Zhu	41.61	—
Yu Zhu	41.81	—
Mai Men Dong	42.40	—
Shan Yu Rou	44.01	—
Zhi Ke	44.80	—
Bai He	45.18	—
Cang Zhu	45.25	—
Wu Wei Zi	45.26	—
Qiang Huo	45.32	—
Jie Geng	45.34	—
Ma Huang	45.56	—
Sang Zhi	46.23	—
Huang Lian	46.43	—
Jin Yin Hua	46.52	—
Huang Qin	48.81	—
Hou Po	49.14	—

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	9.27
62597	62597	0.00	9.20
62598	62598	0.00	9.44
62430	62430	0.00	8.80

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 2.1-2020-01

Validated substance/substance group **Fu Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60004-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Fu Shen; Poriae cocos sclerotium paradicis

Special notes

When selecting the *Fu Shen* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 849 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	1	0	4	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.8837 % (> 99.2960 %)	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
Yi Yi Ren	6.89	—
Ma Huang Gen	7.63	—
Lian Zi	10.05	—
Zhu Ru	11.26	—
Rou Gui	11.53	—
Sheng Jiang	13.78	—
Fu Xiao Mai	13.89	—
Fu Ling	14.10	—
Gui Zhi	14.10	—
Ban Xia (Jiang)	15.00	—
Ji Xue Teng	18.29	—
Gou Teng	19.56	—
Di Gu Pi	23.17	—
Long Yan Rou	24.62	—
Bai Bian Dou	25.33	—
Bai Xian Pi	25.68	—
He Huan Pi	26.28	—
Bai Shao Yao	26.56	—
Bai Zi Ren	27.26	—
Ji Li	27.56	—
Ling Zhi	28.67	—
Tao Ren	29.20	—
Tian Men Dong	29.49	—
Fo Shou	29.83	—
Da Zao	29.83	—
Ci Wu Jia	31.13	—
Ye Jiao Teng	31.69	—
Shen Qu	31.73	—
Huo Ma Ren	32.07	—
Wang Bu Liu Xing	32.52	—
Chen Pi	32.53	—
San Qi	32.77	—
Tai Zi Shen	33.69	—
Zhi Mu	33.88	—
Gua Lou	37.17	—
Zhe Bei Mu	37.27	—
Mang Xiao	37.83	—
Yuan Zhi	38.62	—
Tu Fu Ling	38.72	—
Ren Dong Teng	39.91	—
Pi Pa Ye	40.50	—
Dan Shen	40.83	—
Yan Hu Suo	40.87	—
Dan Dou Chi	41.19	—
Chuan Xiong	41.67	—
Tian Hua Fen	42.20	—
Dang Shen	42.92	—
Yu Jin	43.31	—
Lian Qiao	43.96	—
Shan Yu Rou	44.43	—
Shan Yao	45.82	—
Zhi Ke	46.53	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	48.46	—
Huang Qin	48.82	—
Mao Dong Qing	48.90	—
Bai Zhu	50.04	—

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	7.36
62618	62618	0.00	6.89

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 2.1-2020-01

Validated substance/substance group **Fu Xiao Mai**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60081-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Fu Xiao Mai; Triticum aestivum 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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G244HS021QM1	62183	40	from supplier
PhytoComm	Fu Xiao Mai	G244HS021QM1	62184	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G244HS021QM1	62183 [†]	20
PhytoComm	Fu Xiao Mai	G244HS021QM1	62184 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	11	225	15	25 550
Type B	20	105	15	13 027
Type C	0	0	0	853

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	99.9709 % (> 99.9560 %)	93.7500 % (> 92.5000 %)
Type B	99.8837 % (> 99.8548 %)	87.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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	2.70	—
Rou Gui	2.84	—
Zhu Ru	3.55	—
Sheng Jiang	6.75	—
Fu Ling	7.31	—
Gui Zhi	7.42	—
Ji Li	7.45	—
Di Gu Pi	8.13	—
Bai Bian Dou	8.17	—
Wang Bu Liu Xing	8.52	—
Ban Xia (Jiang)	8.80	—
He Huan Pi	8.88	—
Ci Wu Jia	9.14	—
Bai Shao Yao	9.26	—
Lian Zi	10.85	—
Ling Zhi	11.20	—
Bai Zi Ren	11.46	—
Fu Shen	11.79	—
Tao Ren	13.02	—
Shen Qu	13.43	—
Ma Huang Gen	13.63	—
Tian Hua Fen	13.69	—
Gou Teng	14.68	—
Huo Ma Ren	14.89	—
Zhe Bei Mu	15.58	—
Ji Xue Teng	15.67	—
Tai Zi Shen	15.91	—
Shan Yao	16.64	—
Gua Lou	17.09	—
Lian Qiao	18.89	—
Bai Xian Pi	18.95	—
Pi Pa Ye	19.06	—
Ren Dong Teng	19.24	—
Yu Jin	19.43	—
Yuan Zhi	19.65	—
Chen Pi	21.37	—
Ye Jiao Teng	22.45	—
Tian Men Dong	23.07	—
Fo Shou	23.54	—
Chuan Xiong	23.98	—
Fu Zi	24.03	—
Long Yan Rou	24.20	—
Yan Hu Suo	25.48	—
Da Zao	25.97	—
Tu Fu Ling	26.37	—
Mao Dong Qing	26.43	—
Dan Shen	26.77	—
Dang Shen	26.99	—
San Qi	27.40	—
Shan Yu Rou	28.90	—
Che Qian Zi	28.99	—
Wu Wei Zi	29.10	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dang Gui Wei	29.12	—
Fang Ji	29.51	—
Jiao Gu Lan	29.81	—
Dan Dou Chi	30.77	—
Chai Hu	30.80	—
Zhi Mu	32.03	—
Jin Yin Hua	32.03	—
Hou Po	32.15	—
Yin Yang Huo	33.22	—
Cang Zhu	34.13	—
Dang Gui	34.22	—
Yu Zhu	34.65	—
Suan Zao Ren	35.93	—
Zhi Ke	36.76	—
Huang Qin	36.80	—
Mang Xiao	38.69	—
Ma Huang	39.92	—
Qing Pi	43.41	—
Qiang Huo	43.88	—
Gou Qi Zi	44.34	—
Hong Jing Tian	44.72	—
Du Zhong	44.83	—
Huang Lian	45.17	—
Ze Xie	46.92	—
Bai Zhu	48.34	—
Sang Zhi	48.66	—
Mu Dan Pi	48.88	—
Jie Geng	49.00	—
Mai Men Dong	49.66	—

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
62631	62631	0.00	4.38
62632	62632	0.00	2.70
62423	62423	0.00	4.20
62184	62184	0.00	2.84
62424	62424	0.00	2.79
62183	62183	0.00	3.10

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 2.1-2020-01

Validated substance/substance group **Fu Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50884-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Fu Zi	3	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	G005H0801624	62081	40	from supplier
PhytoComm	Fu Zi	G005H0801624	62082	40	from supplier
PhytoComm	Fu Zi	G005H0801723	62225	40	from supplier
PhytoComm	Fu Zi	G005H0801723	62226	40	from supplier
PhytoComm	Fu Zi	G005HS158SP1	62699	40	from supplier
PhytoComm	Fu Zi	G005HS158SP1	62700	40	from supplier

Validation samples

A total of 39 821 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 Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 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	G005H0801624	62081 [†]	20
PhytoComm	Fu Zi	G005H0801624	62082 [†]	20
PhytoComm	Fu Zi	G005H0801723	62225 [†]	20
PhytoComm	Fu Zi	G005H0801723	62226 [†]	20
PhytoComm	Fu Zi	G005HS158SP1	62699 [†]	20
PhytoComm	Fu Zi	G005HS158SP1	62700 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	g005h0801322	1
Phytocomm	Fu Zi	G005H0801322	1
Phytocomm	Fu Zi	G005H0801524	1
Phytocomm	Fu Zi	g005h0801122	1

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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	238	2	25 561
Type B	0	107	13	13 047
Type C	1	0	4	848

The substance/substance group *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	100.0000 % (> 99.9701 %)	99.1667 % (> 97.9167 %)
Type B	100.0000 % (> 99.9421 %)	89.1667 % (> 86.6667 %)
Type C	99.8837 % (> 99.2960 %)	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 *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	3.94	—
Pi Pa Ye	4.62	—
Gui Zhi	6.14	—
Ren Dong Teng	7.07	—
Lian Qiao	7.90	—
Bai Bian Dou	8.35	—
Ji Li	8.74	—
Hou Po	9.49	—
Tian Hua Fen	9.67	—
Jin Yin Hua	9.95	—
Chuan Xiong	10.41	—
Bai Zi Ren	10.56	—
Gua Lou	10.61	—
Shen Qu	10.78	—
Huo Ma Ren	11.00	—
Jiang Huang	12.09	—
Yan Hu Suo	13.23	—
Bai Shao Yao	13.45	—
Ji Xue Teng	14.10	—
Jiao Gu Lan	14.20	—
Tai Zi Shen	14.54	—
Lian Zi	14.86	—
Bu Gu Zhi	15.04	—
He Huan Pi	15.30	—
Yuan Zhi	15.34	—
Cang Zhu	15.36	—
Che Qian Zi	15.36	—
Sha Ren	15.53	—
Dan Shen	15.58	—
Ling Zhi	15.71	—
Zhe Bei Mu	15.99	—
Yin Yang Huo	16.37	—
Chen Pi	16.78	—
Fang Ji	16.82	—
Fu Ling	16.90	—
Ban Xia (Jiang)	17.12	—
Gu Sui Bu	17.21	—
Ma Bian Cao	17.43	—
San Qi	17.88	—
Chai Hu	17.90	—
Mao Dong Qing	18.00	—
Wu Yao	18.25	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	18.40	—
Tao Ren	18.46	—
Wu Zhu Yu	18.48	—
Hong Jing Tian	18.52	—
Wu Wei Zi	18.59	—
Gou Teng	18.60	—
Dan Dou Chi	18.70	—
Bai Xian Pi	18.82	—
Yu Jin	18.82	—
Fu Xiao Mai	18.98	—
Ye Jiao Teng	19.66	—
Ma Huang	19.69	—
Sheng Jiang	19.72	—
Rou Gui	19.85	—
Qing Pi	20.00	—
Hu Zhang	20.06	—
Qiang Huo	20.75	—
Ci Wu Jia	20.99	—
Di Gu Pi	21.08	—
Zhu Ru	21.50	—
Xin Yi	21.97	—
Shan Yu Rou	21.97	—
Gou Qi Zi	22.02	—
Mu Dan Pi	22.04	—
Da Zao	22.59	—
Sang Ye	23.23	—
Yu Zhu	23.61	—
Yi Mu Cao	23.63	—
Dang Gui	23.75	—
Huang Bo	23.90	—
Dan Zhu Ye	23.95	—
Dang Gui Wei	24.29	—
Xian He Cao	24.55	—
(Bai) Dou Kou	24.59	—
Cang Er Zi	24.64	—
Yi Yi Ren	24.79	—
Huang Lian	24.86	—
Dang Shen	25.18	—
Lu Gen	25.36	—
(Shi) Chang Pu	25.60	—
Ma Huang Gen	25.64	—
Pu Gong Ying	25.88	—
Suan Zao Ren	26.08	—
Xiang Fu	26.24	—
Xi Xian Cao	26.98	—
Ze Xie	27.37	—
Xu Duan	27.66	—
Zi Hua Di Ding	27.86	—
Zhi Ke	28.04	—
Xia Ku Cao	28.37	—
Yin Chen Hao	28.62	—
Du Zhong	28.93	—
Fo Shou	28.97	—
Sang Bai Pi	29.12	—
Chuang Mu Xiang	29.69	—
Zhi Mu	29.98	—
She Gan	30.05	—
Huang Qin	30.17	—
Gan Cao	31.34	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Hua She She Cao	31.44	—
He Shou Wu	32.22	—
Sang Zhi	32.45	—
Long Yan Rou	32.57	—
Han Lian Cao	32.61	—
Nü Zhen Zi	32.65	—
Wei Ling Xian	32.97	—
Gan Jiang	33.16	—
Mang Xiao	33.37	—
Bai Zhu	33.52	—
Ju Hua	33.62	—
Zi Wan	34.03	—
Sang Ji Shend	34.51	—
Wang Bu Liu Xing	35.47	—
Tian Men Dong	35.85	—
Du Huo	35.92	—
E Zhu	36.13	—
Xie Bai	38.25	—
Chi Shao (Yao)	38.47	—
Fu Shen	38.80	—
Fang Feng	39.80	—
Zi Su Zi	40.68	—
Jie Geng	40.76	—
Jing Jie	41.34	—
Ku Shen	41.58	—
Bai Mao Gen	42.32	—
Huang Bai	42.35	—
Zhi Shi	42.64	—
Niu Bang Zi	43.69	—
Huang Jing	44.01	—
Ba Ji Tian	45.54	—
San Leng	46.33	—
Ge Gen	46.38	—
Hong Hua	47.99	—
Huang Qi	48.67	—
Chuan Lian Zi	49.27	—
Wu Mei	49.31	—

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	4.07
62226	62226	0.00	9.37
62225	62225	0.00	9.99
62081	62081	0.00	6.67

continued on the next page

continued from previous page

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
62082	62082	0.00	6.14
62700	62700	0.00	3.94

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 2.1-2020-01

Validated substance/substance group **Gan Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60011-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Gan Cao	2	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

Validation samples

A total of 39 821 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 *Gan Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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 4 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	G119H0521522	8
PhytoComm	Gan Cao	G119H0521121	4
PhytoComm	Gan Cao	G119H0521524	1

- 838 spectra from 12 *Apo-Ident* customers from a total of 512 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	9	6	838

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8222 %)	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 *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
Lian Qiao	18.89	—
Chi Shao (Yao)	19.50	—
Zi Wan	20.67	—
Ze Xie	22.00	—
(Shi) Chang Pu	23.06	—
Zhi Ke	23.20	—
Mu Dan Pi	25.01	—
Yuan Zhi	25.26	—
Shan Yu Rou	25.30	—
Dang Gui Wei	25.62	—
Xiang Fu	27.04	—
Zi Hua Di Ding	27.59	—
San Leng	27.65	—
Dang Gui	28.34	—
Gou Qi Zi	28.43	—
Bai Zhu	28.46	—
Bai Shao Yao	29.20	—
Chuan Xiong	29.33	—
Huang Qin	30.66	—
Niu Bang Zi	30.75	—
Jiao Gu Lan	30.81	—
Wu Wei Zi	31.10	—
Di Gu Pi	31.64	—
Hong Jing Tian	32.07	—
Huang Bai	32.79	—
Chuang Mu Xiang	34.18	—
Jin Yin Hua	34.30	—
Hong Hua	34.47	—
Wu Yao	34.65	—
Jie Geng	35.02	—
Gui Zhi	35.06	—
Chen Pi	35.69	—
Huang Qi	36.21	—
Yin Yang Huo	36.28	—
Hou Po	36.56	—
Ma Huang	36.66	—
Sang Zhi	37.24	—
Shan Yao	37.33	—
Gua Lou	37.39	—
(Huai) Niu Xi	37.76	—
Yu Zhu	37.82	—
Yi Mu Cao	37.94	—
Cang Zhu	37.97	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Mu	38.98	—
He Huan Pi	39.69	—
Zhe Bei Mu	40.69	—
Yan Hu Suo	40.73	—
Ye Jiao Teng	40.91	—
Qing Pi	40.95	—
Fu Zi	41.43	—
San Qi	41.48	—
Tai Zi Shen	41.96	—
Mang Xiao	43.60	—
Pu Gong Ying	43.89	—
Gou Teng	43.97	—
Fang Ji	43.97	—
Dan Shen	43.98	—
She Gan	44.03	—
Chai Hu	44.30	—
Cang Er Zi	44.57	—
Huang Lian	45.28	—
Mu Gua	45.59	—
Bai Mao Gen	45.64	—
Tao Ren	45.96	—
Dang Shen	46.34	—
Pi Pa Ye	47.41	—
Ku Shen	47.77	—
Wei Ling Xian	49.08	—
Qiang Huo	50.07	—

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
62533	62533	0.00	20.16
62341	62341	0.00	20.13
62534	62534	0.00	18.89
62342	62342	0.00	19.50

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 2.1-2020-01

Validated substance/substance group **Gan Jiang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60077-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G252H1145821	62325	40	from supplier
PhytoComm	Gan Jiang	G252H1145821	62326	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G252H1145821	62325 [†]	20
PhytoComm	Gan Jiang	G252H1145821	62326 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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	G252H1145221	1
Phytocomm	Gan Jiang	g252h01145221	1
Phytocomm	Gan Jiang	G252H1145421	1
Phytocomm	Gan Jiang	G252H1145521	1

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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	25 721
Type B	0	40	0	13 127
Type C	2	1	3	847

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.4186 % (> 98.8309 %)	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
Gua Lou	11.46	—
Sang Zhi	11.72	—
Gui Zhi	12.82	—
Cang Er Zi	13.26	—
Huo Ma Ren	14.56	—
Sha Shen (Bei)	15.32	—
Tian Hua Fen	16.08	—
Di Gu Pi	16.58	—
Zi Su Zi	16.81	—
Chai Hu	16.87	—
Sheng Jiang	17.16	—
Xiang Fu	17.27	—
Gou Qi Zi	17.67	—
Yan Hu Suo	18.27	—
(Shi) Chang Pu	19.05	—
Zhi Mu	19.16	—
Dang Shen	19.59	—
Ji Li	19.61	—
Chuan Lian Zi	19.84	—
Jie Geng	20.25	—
She Gan	20.71	—
Niu Bang Zi	20.87	—
(Huai) Niu Xi	21.33	—
San Qi	21.35	—
Dang Gui Wei	21.54	—
Ba Ji Tian	21.57	—
Qing Pi	21.60	—
Jiang Huang	21.83	—
Yi Mu Cao	22.03	—
Zi Wan	22.35	—
Long Dan (Cao)	22.52	—
Xie Bai	23.48	—
Dang Gui	23.98	—
Cang Zhu	24.18	—
E Zhu	24.55	—
Hou Po	24.61	—
Lian Qiao	24.84	—
Lu Gen	25.47	—
Du Huo	25.64	—
Mang Xiao	25.86	—
Lai Fu Zi	26.59	—
Fang Feng	27.01	—
(Bai) Dou Kou	27.34	—
Shan Yao	27.40	—
Chuan Xiong	27.43	—
Sha Ren	28.55	—
Ku Shen	29.10	—
Mu Gua	29.51	—
Bai Mao Gen	29.63	—
Chen Pi	30.47	—
Huang Qi	30.85	—
Wu Yao	32.65	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai He	34.03	—
Bai Zhu	35.19	—
Mu Dan Pi	35.31	—
San Leng	35.39	—
Xian He Cao	35.51	—
Mai Ya	36.39	—
Hong Jing Tian	36.57	—
Bai Bian Dou	37.36	—
Qin Jiao	38.35	—
Mai Men Dong	38.70	—
Tu Fu Ling	39.12	—
Huang Lian	39.69	—
Yuan Zhi	39.81	—
Bai Shao Yao	40.08	—
Ze Xie	40.44	—
Fu Zi	40.45	—
Wei Ling Xian	40.48	—
Yu Jin	40.80	—
Pu Gong Ying	41.17	—
He Huan Pi	41.54	—
Qiang Huo	42.13	—
Dong Gua Zi	42.55	—
Zhi Zi	43.37	—
Wu Wei Zi	43.69	—
Ban Xia (Jiang)	44.14	—
Wang Bu Liu Xing	44.60	—
Zhi Ke	44.88	—
Da Zao	44.96	—
Nü Zhen Zi	45.13	—
Fu Ling	45.27	—
Mo Yao	45.78	—
Jin Yin Hua	45.88	—
Zhu Ling	45.92	—
Sang Ji Shend	46.00	—
Ban Zhi Lian	46.10	—
Huang Bo	46.31	—
Fang Ji	46.31	—
Lian Zi	46.50	—
Xin Yi	46.57	—
Gou Teng	48.10	—
Chi Shao (Yao)	48.16	—
Ju Hua	48.51	—
Ren Dong Teng	49.63	—
Sang Bai Pi	50.19	—

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
62325	62325	0.00	11.46
62326	62326	0.00	11.61

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 2.1-2020-01

Validated substance/substance group **Ge Gen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60050-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ge Gen	2	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	G206H1301621	61949	40	from supplier
PhytoComm	Ge Gen	G206H1301621	61950	40	from supplier
PhytoComm	Ge Gen	G206H1301822	62489	40	from supplier
PhytoComm	Ge Gen	G206H1301822	62490	40	from supplier

Validation samples

A total of 39 821 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 *Ge Gen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G206H1301621	61949 [†]	20
PhytoComm	Ge Gen	G206H1301621	61950 [†]	20
PhytoComm	Ge Gen	G206H1301822	62489 [†]	20
PhytoComm	Ge Gen	G206H1301822	62490 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G206H1301422	2
PhytoComm	Ge Gen	G206H1301323	1

- 848 spectra from 12 *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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	5	848

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8240 %)	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 *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
Sang Ji Shend	22.94	—
Zhi Shi	25.34	—
Hong Jing Tian	34.88	—
Dan Shen	36.07	—
Sang Bai Pi	39.23	—
Jing Jie	39.60	—
Gu Sui Bu	40.29	—
Xu Duan	40.41	—
Dan Zhu Ye	42.36	—
He Shou Wu	42.51	—
Hou Po	43.19	—
Huang Lian	43.92	—
Gou Teng	44.01	—
Xian He Cao	44.06	—
Yin Yang Huo	44.57	—
Pu Gong Ying	45.30	—
Qiang Huo	46.34	—
Hua Shi	49.29	—
He Huan Pi	50.04	—

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
61950	61950	0.00	24.62
62490	62490	0.00	25.34
62489	62489	0.00	24.63
61949	61949	0.00	22.94

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 2.1-2020-01

Validated substance/substance group **Gou Qi Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60094-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G150H0913821	62349	40	from supplier
PhytoComm	Gou Qi Zi	G150H0913821	62350	40	from supplier
PhytoComm	Gou Qi Zi	G150HS171SK1	62615	40	from supplier
PhytoComm	Gou Qi Zi	G150HS171SK1	62616	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G150H0913821	62349 [†]	20
PhytoComm	Gou Qi Zi	G150H0913821	62350 [†]	20
PhytoComm	Gou Qi Zi	G150HS171SK1	62615 [†]	20
PhytoComm	Gou Qi Zi	G150HS171SK1	62616 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G150H0913422	4
PhytoComm	Gou Qi Zi	G150H0913221	1

- 847 spectra from 12 *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 *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	25 641
Type B	0	79	1	13 087
Type C	8	3	3	839

The substance/substance group *Gou Qi 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	98.7500 % (> 95.0000 %)
Type C	98.7508 % (> 98.1626 %)	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 Wei	8.49	—
She Gan	9.26	—
Chen Pi	9.66	—
Chai Hu	9.70	—
Du Huo	9.94	—
Yu Zhu	10.45	—
Bai Mao Gen	11.31	—
Hong Jing Tian	12.96	—
Long Yan Rou	13.19	—
Chuan Xiong	13.47	—
Zi Wan	13.56	—
Gua Lou	13.74	—
Zhi Mu	14.07	—
Cang Zhu	14.14	—
Dang Gui	14.91	—
Zhi Ke	16.33	—
Shan Yu Rou	16.41	—
Gui Zhi	16.50	—
(Huai) Niu Xi	16.63	—
Ma Huang	16.70	—
Huang Qi	16.77	—
Chuang Mu Xiang	16.81	—
Mu Dan Pi	17.58	—
San Qi	17.59	—
Lu Gen	18.03	—
Jin Yin Hua	18.08	—
Sang Zhi	18.33	—
(Shi) Chang Pu	18.61	—
San Leng	19.03	—
Di Gu Pi	19.13	—
Yuan Zhi	19.14	—
Lian Qiao	19.19	—
Jie Geng	19.20	—
Fang Feng	19.55	—
Gan Cao	19.70	—
Qing Pi	19.78	—
Hou Po	20.19	—
Tai Zi Shen	20.23	—
Pi Pa Ye	20.53	—
Jiao Gu Lan	20.72	—
Dang Shen	20.99	—
Ye Jiao Teng	21.54	—
Ren Dong Teng	21.59	—
Tao Ren	21.97	—
He Huan Pi	22.21	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhe Bei Mu	22.71	—
Da Zao	22.79	—
Yin Yang Huo	22.86	—
Huang Qin	22.93	—
Dan Shen	23.01	—
Bai Shao Yao	23.45	—
Zi Hua Di Ding	23.67	—
Ju Hua	23.74	—
Long Dan (Cao)	23.90	—
Gan Jiang	24.22	—
Shan Yao	24.32	—
Fu Zi	24.58	—
Wu Wei Zi	24.73	—
Mu Gua	24.85	—
Yan Hu Suo	25.37	—
Bai Zhu	25.41	—
Ji Xue Teng	26.60	—
Ba Ji Tian	26.77	—
Xiang Fu	26.80	—
Ci Wu Jia	26.82	—
Lian Zi	27.76	—
Bai Bian Dou	27.79	—
Gou Teng	27.83	—
Ji Li	28.29	—
Qin Jiao	28.88	—
Wu Yao	29.27	—
Shen Qu	29.27	—
Xie Bai	29.74	—
Bai Xian Pi	29.88	—
Yi Mu Cao	29.99	—
Bai Zi Ren	29.99	—
Mang Xiao	30.03	—
Tian Hua Fen	30.57	—
Ze Xie	30.62	—
Tu Fu Ling	30.66	—
Fang Ji	31.22	—
Fo Shou	31.43	—
Chi Shao (Yao)	31.75	—
Ban Xia (Jiang)	32.12	—
Rou Gui	32.53	—
Huo Ma Ren	32.69	—
Xian He Cao	32.75	—
Sheng Jiang	33.84	—
Lai Fu Zi	34.69	—
Mao Dong Qing	34.78	—
E Zhu	34.84	—
Dan Dou Chi	35.42	—
Che Qian Zi	35.64	—
Jiang Huang	35.74	—
Ku Shen	35.82	—
Suan Zao Ren	35.83	—
Wei Ling Xian	37.22	—
Tian Men Dong	37.28	—
Fu Ling	37.61	—
Sha Shen (Bei)	37.64	—
Qiang Huo	38.38	—
Niu Bang Zi	38.79	—
Huang Lian	38.97	—
Ling Zhi	39.04	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Hong Hua	39.77	—
Chuan Lian Zi	40.15	—
Zi Su Zi	40.61	—
Sha Ren	40.71	—
Cang Er Zi	40.93	—
Fu Xiao Mai	41.45	—
Huang Bai	41.90	—
Zhu Ru	41.98	—
Mai Men Dong	41.98	—
Bai He	42.14	—
Fu Shen	43.62	—
Yu Jin	43.80	—
Yi Yi Ren	44.33	—
Gu Sui Bu	44.42	—
Zhi Zi	46.58	—
(Bai) Dou Kou	47.32	—
Huang Bo	47.66	—
Du Zhong	47.95	—
Ma Huang Gen	48.79	—
Huang Jing	48.98	—
Ma Bian Cao	49.11	—
Ge Gen	49.36	—
Sang Ye	49.46	—

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
62616	62616	0.00	10.14
62349	62349	0.00	8.49
62350	62350	0.00	9.05
62615	62615	0.00	9.66

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 2.1-2020-01

Validated substance/substance group **Gou Teng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60097-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 850 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 *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	0	159	1	25 641
Type B	0	76	4	13 087
Type C	2	0	3	848

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	100.0000 % (> 99.9701 %)	99.3750 % (> 97.5000 %)
Type B	100.0000 % (> 99.9422 %)	95.0000 % (> 91.2500 %)
Type C	99.5349 % (> 98.9477 %)	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	3.01	—
Gui Zhi	5.68	—
Ma Huang Gen	6.54	—
Yi Yi Ren	6.63	—
Huo Ma Ren	8.00	—
Ji Xue Teng	8.00	—
Sheng Jiang	8.67	—
Lian Zi	8.78	—
Fu Ling	8.83	—
Ban Xia (Jiang)	9.82	—
He Huan Pi	9.96	—
Fu Shen	10.09	—
Zhu Ru	10.22	—
Ci Wu Jia	10.94	—
Tai Zi Shen	11.38	—
Tao Ren	12.25	—
Fu Xiao Mai	12.46	—
Bai Bian Dou	12.88	—
Bai Shao Yao	13.01	—
Ren Dong Teng	13.22	—
Bai Xian Pi	13.87	—
Ling Zhi	14.02	—
Di Gu Pi	15.49	—
Ji Li	16.32	—
Zhe Bei Mu	16.67	—
Fo Shou	17.37	—
Ye Jiao Teng	17.51	—
Shen Qu	18.08	—
Bai Zi Ren	18.75	—
Huang Bo	19.68	—
Tian Men Dong	20.50	—
Tian Hua Fen	20.68	—
Dan Dou Chi	20.79	—
Chen Pi	21.86	—
Yu Jin	22.08	—
Yan Hu Suo	22.11	—
Wang Bu Liu Xing	22.41	—
Nü Zhen Zi	22.44	—
Dan Shen	23.22	—
Da Zao	23.44	—
Dan Zhu Ye	23.63	—
Chuan Xiong	23.65	—
Yi Mu Cao	23.69	—
Tu Fu Ling	24.27	—
Lian Qiao	25.06	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jing Jie	25.23	—
Long Yan Rou	25.44	—
Yin Chen Hao	25.71	—
Gua Lou	26.11	—
Mao Dong Qing	26.52	—
Qing Hao	26.95	—
Qiang Huo	27.17	—
Yin Yang Huo	27.53	—
Huang Lian	27.66	—
Pi Pa Ye	27.79	—
Xu Duan	28.04	—
Pu Gong Ying	28.65	—
Shan Yao	28.73	—
Yuan Zhi	29.56	—
Ma Huang	29.62	—
Hong Jing Tian	30.56	—
Chai Hu	30.93	—
San Qi	31.00	—
Jiao Gu Lan	31.35	—
Fu Zi	32.56	—
Fang Ji	33.24	—
Wu Zhu Yu	33.33	—
Yu Zhu	34.80	—
Zhi Mu	34.84	—
Mang Xiao	35.54	—
Zhi Ke	35.68	—
Jin Yin Hua	36.66	—
Dang Shen	37.38	—
Wu Wei Zi	37.52	—
Dang Gui Wei	38.41	—
Cang Zhu	38.50	—
Che Qian Zi	38.67	—
Xian He Cao	38.81	—
Shan Yu Rou	39.27	—
Hou Po	39.27	—
Sang Ji Shend	40.60	—
Bu Gu Zhi	40.66	—
Huang Qin	41.37	—
Ma Bian Cao	41.41	—
Sha Ren	41.88	—
Zhi Shi	42.02	—
Xin Yi	42.86	—
Hu Zhang	43.11	—
Sang Zhi	43.92	—
Wei Ling Xian	44.24	—
Jie Geng	44.94	—
Xuan Shen	45.89	—
Rou Cong Rong	46.34	—
Dang Gui	46.36	—
Bai Zhu	46.63	—
Xia Ku Cao	47.14	—
Wu Jia Pi	47.27	—
Jin Qian Cao	47.58	—
Du Zhong	47.59	—
Han Lian Cao	47.85	—
Bai He	48.35	—
Qing Pi	49.84	—
(Sheng) Di Huang	50.81	—

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
62630	62630	0.00	3.01
62314	62314	0.00	20.75
62313	62313	0.00	19.68
62629	62629	0.00	5.68

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 2.1-2020-01

Validated substance/substance group **Gu Sui Bu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002263-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Gu Sui Bu	1	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

Validation samples

A total of 39 821 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 *Gu Sui Bu*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	G094H1011221	2
Phytocomm	Gu Sui Bu	g094h1011023	1
Phytocomm	Gu Sui Bu	g094h1011221	1
PhytoComm	Gu Sui Bu	G094H1011221	1
PhytoComm	Gu Sui Bu	G094H1011522	1

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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	80	0	25 721
Type B	0	40	0	13 127
Type C	3	0	6	844

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.4574 % (> 98.8691 %)	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 *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
Ma Bian Cao	8.76	—
He Huan Pi	14.46	—
Xian He Cao	16.39	—
Bu Gu Zhi	17.37	—
He Shou Wu	17.43	—
Yu Jin	18.16	—
Gui Zhi	18.58	—
Sang Bai Pi	18.60	—
Fu Zi	19.64	—
Yin Yang Huo	19.67	—
Xu Duan	21.30	—
Chen Pi	22.26	—
Yi Mu Cao	22.43	—
Dan Zhu Ye	22.48	—
Wu Yao	22.75	—
Jiang Huang	23.74	—
Sang Ji Shend	24.74	—
Ji Li	25.92	—
Huang Bo	29.31	—
Nü Zhen Zi	29.88	—
Fang Feng	29.93	—
Sha Ren	29.94	—
Dan Shen	30.16	—
Yin Chen Hao	30.37	—
Mang Xiao	30.53	—
Wu Zhu Yu	30.69	—
(Bai) Dou Kou	30.82	—
Lu Gen	31.21	—
Hu Zhang	31.95	—
Huang Jing	33.41	—
Qiang Huo	34.35	—
E Zhu	36.31	—
Xia Ku Cao	37.70	—
Hong Jing Tian	37.81	—
Shan Yao	38.56	—
Jing Jie	38.63	—
Ge Gen	39.21	—
Cang Er Zi	40.39	—
Hou Po	41.94	—
Huang Lian	42.27	—
Ju Hua	42.43	—
Gou Qi Zi	42.71	—
Xiang Fu	42.89	—
Xi Xian Cao	42.93	—
Chuan Xiong	43.59	—
Lian Qiao	44.90	—
Qing Pi	45.06	—
Zi Su Zi	45.07	—
(Sheng) Di Huang	45.58	—
Gan Jiang	45.99	—
Mu Dan Pi	46.28	—
Gua Lou	47.25	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xin Yi	47.72	—
Sang Zhi	47.90	—
She Gan	48.16	—
(Shi) Chang Pu	48.78	—
Gou Teng	49.06	—
Bai Hua She She Cao	49.23	—
Tu Fu Ling	49.74	—
Chai Hu	49.83	—
Hua Shi	50.30	—

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
62780	62780	0.00	8.76
62779	62779	0.00	9.10

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 2.1-2020-01

Validated substance/substance group **Gua Lou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60128-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G241H1096621	61913	40	from supplier
PhytoComm	Gua Lou	G241H1096621	61914	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G241H1096621	61913 [†]	20
PhytoComm	Gua Lou	G241H1096621	61914 [†]	20
PhytoComm	Gua Lou	G241H1096821	62317 [†]	20
PhytoComm	Gua Lou	G241H1096821	62318 [†]	20
PhytoComm	Gua Lou	G241HS202SK1	62547 [†]	20
PhytoComm	Gua Lou	G241HS202SK1	62548 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 850 spectra from 12 *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 *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	25 560
Type B	1	120	0	13 046
Type C	1	3	0	849

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	99.9927 % (> 99.9778 %)	99.5833 % (> 98.3333 %)
Type B	99.9855 % (> 99.9565 %)	100.0000 % (> 95.0000 %)
Type C	99.8450 % (> 99.2578 %)	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
Fo Shou	4.53	—
Chen Pi	7.40	—
Bai Shao Yao	7.90	—
Ye Jiao Teng	8.63	—
Dang Gui Wei	9.55	—
Tian Men Dong	10.17	—
Tai Zi Shen	10.20	—
Tao Ren	10.52	—
Shen Qu	10.59	—
Gan Jiang	10.75	—
Yuan Zhi	10.98	—
Zhe Bei Mu	11.14	—
Dang Shen	11.27	—
Di Gu Pi	11.30	—
He Huan Pi	11.31	—
Pi Pa Ye	12.06	—
Gou Qi Zi	12.53	—
Rou Gui	12.75	—
Yu Zhu	12.93	—
Ci Wu Jia	12.99	—
(Huai) Niu Xi	13.02	—
Sang Zhi	13.28	—
Da Zao	13.90	—
Zhi Mu	14.06	—
Huo Ma Ren	14.31	—
Cang Zhu	14.51	—
Ban Xia (Jiang)	14.51	—
Lian Zi	14.57	—
She Gan	14.71	—
Xie Bai	14.79	—
Chai Hu	15.02	—
(Shi) Chang Pu	15.08	—
Ji Li	15.19	—
Chuan Xiong	15.21	—
Bai Zi Ren	15.49	—
Bai Bian Dou	16.07	—
Gui Zhi	16.21	—
Ku Shen	16.52	—
Qing Pi	16.55	—
Tu Fu Ling	16.88	—
San Qi	16.95	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Teng	17.71	—
Ren Dong Teng	17.99	—
Shan Yao	17.99	—
Sheng Jiang	18.18	—
Niu Bang Zi	18.27	—
Zi Wan	18.41	—
Lian Qiao	18.52	—
Fu Ling	18.58	—
Jie Geng	19.17	—
Wei Ling Xian	19.38	—
Xiang Fu	19.38	—
Ji Xue Teng	20.26	—
Tian Hua Fen	20.29	—
Dang Gui	20.48	—
Bai Xian Pi	20.75	—
Zi Su Zi	21.03	—
Zhu Ru	21.10	—
Ma Huang	21.11	—
Fang Feng	21.26	—
Fu Xiao Mai	21.32	—
Long Yan Rou	21.87	—
Yi Mu Cao	22.02	—
Shan Yu Rou	22.64	—
Yan Hu Suo	23.11	—
Ling Zhi	23.20	—
Hou Po	23.22	—
Fu Zi	23.55	—
Dan Shen	23.66	—
Bai Mao Gen	23.99	—
Du Huo	24.38	—
Huang Qi	24.70	—
Yi Yi Ren	24.98	—
Jiao Gu Lan	25.15	—
Zhi Ke	25.51	—
Jiang Huang	26.07	—
Bai Zhu	26.42	—
Yin Yang Huo	26.56	—
Ba Ji Tian	26.92	—
Cang Er Zi	26.95	—
Ze Xie	27.19	—
Mu Gua	27.33	—
Huang Qin	27.34	—
Mang Xiao	27.68	—
Ma Huang Gen	27.73	—
Mao Dong Qing	27.86	—
Wu Yao	29.06	—
Jin Yin Hua	29.15	—
Zhi Zi	29.30	—
(Bai) Dou Kou	29.40	—
E Zhu	29.80	—
Lu Gen	29.91	—
Wu Wei Zi	29.98	—
Hong Jing Tian	30.01	—
Chuan Lian Zi	30.57	—
Fu Shen	30.97	—
San Leng	31.16	—
Huang Bo	31.82	—
Lai Fu Zi	32.05	—
Xu Duan	32.12	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Shen (Bei)	32.60	—
Sha Ren	32.70	—
Dan Dou Chi	32.73	—
Nü Zhen Zi	32.99	—
Fang Ji	33.43	—
Yu Jin	33.62	—
Gan Cao	33.66	—
Chuang Mu Xiang	33.73	—
Long Dan (Cao)	33.99	—
Qiang Huo	34.61	—
Mu Dan Pi	34.66	—
Xian He Cao	34.72	—
Che Qian Zi	34.80	—
Ju Hua	36.42	—
Dong Gua Zi	37.18	—
Wang Bu Liu Xing	37.61	—
Ma Bian Cao	37.80	—
Suan Zao Ren	38.00	—
Bai He	39.03	—
Pu Gong Ying	39.09	—
Qin Jiao	39.46	—
Chi Shao (Yao)	39.50	—
Xin Yi	41.28	—
Huang Lian	41.89	—
Wu Zhu Yu	42.23	—
Mai Men Dong	42.29	—
Bu Gu Zhi	46.83	—
Gu Sui Bu	48.13	—
Sang Ji Shend	49.09	—
Sang Bai Pi	49.35	—
Xi Xian Cao	49.77	—
Mai Ya	49.82	—

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
61914	61914	0.00	10.75
61913	61913	0.00	11.41
62317	62317	0.00	14.97
62547	62547	0.00	5.42
62318	62318	0.00	14.31
62548	62548	0.00	4.53

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 2.1-2020-01

Validated substance/substance group	Guang Huo Xiang
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	12/03/2020
Report number	61079-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Guang Huo Xiang	1	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

Validation samples

A total of 39 821 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 *Guang Huo Xiang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	0	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

The substance/substance group *Guang Huo 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Ai Ye	22.05	—
Xuan Shen	38.86	—
Jue Ming Zi	39.90	—
Ye Jiao Teng	39.96	—
Pu Gong Ying	42.57	—
Hua Shi	44.46	—
Shu Di (Huang)	47.43	—
Wu Mei	48.37	—

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	23.35
62590	62590	0.00	22.05

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 2.1-2020-01

Validated substance/substance group	Gui Zhi
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	12/03/2020
Report number	60189-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Gui Zhi	5	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	G070H1001621	61777	40	from supplier
PhytoComm	Gui Zhi	G070H1001621	61815	40	from supplier
PhytoComm	Gui Zhi	G070H1001621	61873	40	from supplier
PhytoComm	Gui Zhi	G070H1001621	61874	40	from supplier
PhytoComm	Gui Zhi	G070H1001621	62065	40	from supplier
PhytoComm	Gui Zhi	G070H1001621	62066	40	from supplier
PhytoComm	Gui Zhi	G070H1001624	62103	40	from supplier
PhytoComm	Gui Zhi	G070H1001624	62104	40	from supplier
PhytoComm	Gui Zhi	G070H1001821	62289	40	from supplier
PhytoComm	Gui Zhi	G070H1001821	62290	40	from supplier
PhytoComm	Gui Zhi	G070H1001822	62415	40	from supplier
PhytoComm	Gui Zhi	G070H1001822	62416	40	from supplier
PhytoComm	Gui Zhi	G070HS205SH1	62563	40	from supplier
PhytoComm	Gui Zhi	G070HS205SH1	62564	40	from supplier

Validation samples

A total of 39 821 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 560 spectra of 14 reference samples from the substance/substance group *Gui Zhi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 5 different batches.
- 25 241 spectra from a total of 314 batches from further 172 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.

- 280 spectra of 14 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	G070H1001621	61777 [†]	20
PhytoComm	Gui Zhi	G070H1001621	61815 [†]	20
PhytoComm	Gui Zhi	G070H1001621	61873 [†]	20
PhytoComm	Gui Zhi	G070H1001621	61874 [†]	20
PhytoComm	Gui Zhi	G070H1001621	62065 [†]	20
PhytoComm	Gui Zhi	G070H1001621	62066 [†]	20
PhytoComm	Gui Zhi	G070H1001624	62103 [†]	20
PhytoComm	Gui Zhi	G070H1001624	62104 [†]	20
PhytoComm	Gui Zhi	G070H1001821	62289 [†]	20
PhytoComm	Gui Zhi	G070H1001821	62290 [†]	20
PhytoComm	Gui Zhi	G070H1001822	62415 [†]	20
PhytoComm	Gui Zhi	G070H1001822	62416 [†]	20
PhytoComm	Gui Zhi	G070HS205SH1	62563 [†]	20
PhytoComm	Gui Zhi	G070HS205SH1	62564 [†]	20

- 12887 spectra from a total of 315 batches from further 172 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

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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

[†]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*.

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	560	0	25 241
Type B	9	278	2	12 878
Type C	6	1	5	841

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	100.0000 % (> 99.9700 %)	100.0000 % (> 98.9286 %)
Type B	99.9564 % (> 99.9274 %)	99.2857 % (> 98.2143 %)
Type C	99.0698 % (> 98.4815 %)	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 *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
Sheng Jiang	4.95	—
Ban Xia (Jiang)	6.58	—
Rou Gui	6.66	—
Gou Teng	6.78	—
Yi Yi Ren	7.47	—
Fu Zi	7.68	—
Lian Zi	8.16	—
Tao Ren	8.30	—
Fu Ling	8.36	—
Zhu Ru	8.55	—
He Huan Pi	8.60	—
Di Gu Pi	8.61	—
Ci Wu Jia	8.99	—
Sang Zhi	9.58	—
Jiang Huang	9.62	—
Fu Xiao Mai	10.94	—
Yi Mu Cao	11.16	—
Tai Zi Shen	11.17	—
Ma Huang Gen	12.02	—
Xiang Fu	12.33	—
Tian Hua Fen	12.50	—
San Leng	12.51	—
Lu Gen	12.71	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Ren	12.88	—
Bai Bian Dou	12.88	—
Ji Li	13.46	—
Ji Xue Teng	13.76	—
Bai Zi Ren	13.78	—
Huo Ma Ren	13.78	—
(Shi) Chang Pu	13.88	—
Bai Xian Pi	14.65	—
Yin Yang Huo	14.74	—
Gan Jiang	14.98	—
Hou Po	15.31	—
Ling Zhi	15.41	—
Bai Shao Yao	15.44	—
She Gan	15.50	—
Du Huo	15.70	—
Shen Qu	15.74	—
Chuan Xiong	15.77	—
Bu Gu Zhi	15.94	—
Dang Gui	15.98	—
Zhe Bei Mu	16.14	—
Gua Lou	16.18	—
Sha Shen (Bei)	16.56	—
Jie Geng	16.66	—
Dang Gui Wei	16.66	—
Fu Shen	16.74	—
Wu Yao	16.99	—
Fang Feng	17.02	—
Tian Men Dong	17.34	—
Chai Hu	17.51	—
Long Dan (Cao)	18.20	—
Ye Jiao Teng	18.43	—
Qing Pi	18.43	—
Long Yan Rou	18.73	—
Chen Pi	18.82	—
Fo Shou	18.86	—
Xian He Cao	19.34	—
Dan Shen	19.48	—
(Huai) Niu Xi	19.52	—
Gou Qi Zi	19.99	—
Chuan Lian Zi	20.20	—
Mu Dan Pi	20.64	—
Ba Ji Tian	20.68	—
Da Zao	20.69	—
Cang Zhu	20.74	—
Bai Mao Gen	20.96	—
Ma Bian Cao	21.30	—
Ren Dong Teng	21.52	—
Tu Fu Ling	21.71	—
Yu Jin	21.75	—
E Zhu	22.21	—
Wang Bu Liu Xing	22.28	—
Huang Qi	23.13	—
Dang Shen	23.44	—
Lian Qiao	23.47	—
Zhi Ke	23.61	—
Zi Wan	23.65	—
Gu Sui Bu	23.89	—
Hu Zhang	23.94	—
Qiang Huo	24.16	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	24.35	—
Dan Zhu Ye	24.45	—
Bai Zhu	24.49	—
Shan Yao	24.85	—
Xin Yi	25.04	—
Yan Hu Suo	25.24	—
(Bai) Dou Kou	25.36	—
Mang Xiao	25.36	—
Hong Jing Tian	25.57	—
Xia Ku Cao	26.03	—
Pi Pa Ye	26.14	—
Huang Lian	26.16	—
Cang Er Zi	26.27	—
Yin Chen Hao	26.37	—
Huang Bo	26.54	—
Wu Zhu Yu	27.60	—
Dan Dou Chi	27.69	—
Zhi Mu	28.03	—
Sang Ye	28.03	—
Lai Fu Zi	28.45	—
Ban Zhi Lian	28.49	—
Ze Xie	28.72	—
Bai He	28.82	—
Qin Jiao	29.09	—
Zi Su Zi	29.43	—
Bai Hua She She Cao	29.94	—
Mao Dong Qing	30.41	—
Pu Gong Ying	30.43	—
Xie Bai	30.99	—
Han Lian Cao	31.32	—
Ma Huang	31.35	—
Ju Hua	31.38	—
Sang Bai Pi	31.54	—
San Qi	31.76	—
Wei Ling Xian	31.79	—
Yu Zhu	31.84	—
Chi Shao (Yao)	31.92	—
Xu Duan	32.09	—
Du Zhong	32.44	—
Nü Zhen Zi	32.63	—
Niu Bang Zi	32.78	—
Xi Xian Cao	33.12	—
He Shou Wu	33.31	—
Jiao Gu Lan	33.37	—
Sang Ji Shend	34.24	—
Shan Yu Rou	34.47	—
Mu Gua	34.77	—
Wu Wei Zi	35.77	—
Mai Ya	36.03	—
Huang Jing	36.06	—
Fang Ji	36.14	—
Ku Shen	36.69	—
Zhu Ling	37.69	—
Huang Qin	37.76	—
Jin Yin Hua	38.30	—
Mai Men Dong	39.11	—
Jing Jie	42.35	—
Ge Gen	44.76	—
Zhi Shi	45.04	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Che Qian Zi	46.36	—
Gan Cao	48.59	—
Dong Gua Zi	48.73	—
Wu Mei	49.14	—
Zhi Zi	49.51	—
Zi Hua Di Ding	49.85	—

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
61873	61873	0.00	9.11
62066	62066	0.00	19.22
62289	62289	0.00	8.61
61874	61874	0.00	7.68
62416	62416	0.00	12.71
62103	62103	0.00	9.79
62290	62290	0.00	9.22
62564	62564	0.00	4.95
61777	61777	0.00	9.77
62104	62104	0.00	9.62
62065	62065	0.00	19.20
62415	62415	0.00	13.08
62563	62563	0.00	5.61
61815	61815	0.00	9.17

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 2.1-2020-01

Validated substance/substance group **Han Lian Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60540-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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	G095H0770521	4
PhytoComm	Han Lian Cao	G095H0770321	2

- 845 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	8	845

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8230 %)	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	6.37	—
Xia Ku Cao	8.14	—
Xi Xian Cao	12.90	—
Pu Gong Ying	14.64	—
Tu Fu Ling	15.99	—
Xin Yi	17.07	—
Jiang Huang	18.08	—
Hu Zhang	20.07	—
Sang Ye	20.16	—
Huang Lian	20.27	—
Shan Yao	21.71	—
Gui Zhi	22.17	—
Nü Zhen Zi	23.70	—
Zhi Shi	24.25	—
Yin Yang Huo	27.05	—
He Huan Pi	27.44	—
Cang Er Zi	28.80	—
Sha Ren	30.21	—
Fu Zi	30.54	—
Rou Cong Rong	31.39	—
Yi Mu Cao	34.19	—
Gu Sui Bu	36.00	—
Shen Qu	36.09	—
Du Zhong	36.88	—
Yin Chen Hao	37.77	—
Yan Hu Suo	39.01	—
Bu Gu Zhi	39.02	—
Qiang Huo	39.44	—
Yu Jin	39.80	—
Xuan Shen	40.31	—
Huang Bo	40.55	—
Mang Xiao	41.54	—
Wei Ling Xian	42.82	—
Dan Zhu Ye	43.35	—
Gou Teng	43.36	—
Qing Pi	43.36	—
Ji Li	44.06	—
E Zhu	44.11	—
Dan Shen	44.19	—
Wu Yao	45.58	—
Fang Ji	46.59	—
Chen Pi	46.82	—
Hou Po	47.22	—
Xiang Fu	47.31	—
Wu Wei Zi	48.44	—
(Bai) Dou Kou	48.64	—
Bai Xian Pi	48.69	—
Zhi Ke	49.00	—
Sang Ji Shend	49.55	—
Sang Bai Pi	50.03	—

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
62680	62680	0.00	9.62
62679	62679	0.00	6.37

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 2.1-2020-01

Validated substance/substance group **He Huan Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60175-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
He Huan Pi	2	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

Validation samples

A total of 39 821 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 Huan Pi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	8	159	1	25 633
Type B	6	77	3	13 081
Type C	0	1	1	851

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:

	Specificity	Recognition rate
Type A	99.9516 % (> 99.9366 %)	99.3750 % (> 97.5000 %)
Type B	99.9201 % (> 99.8911 %)	96.2500 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Ban Xia (Jiang)	3.40	—
Tao Ren	4.05	—
Ci Wu Jia	4.88	—
Sheng Jiang	5.39	—
Gui Zhi	6.09	—
Di Gu Pi	7.32	—
Lian Zi	7.43	—
Rou Gui	7.62	—
Bai Zi Ren	7.80	—
Bai Bian Dou	8.50	—
Huo Ma Ren	8.94	—
Bai Shao Yao	9.02	—
Ji Li	10.30	—
Tai Zi Shen	10.69	—
Gou Teng	11.11	—
Shen Qu	11.39	—
Zhe Bei Mu	11.46	—
Yi Yi Ren	11.74	—
Fu Xiao Mai	11.94	—
Ling Zhi	12.26	—
Zhu Ru	12.41	—
Tian Hua Fen	12.69	—
Bai Xian Pi	12.86	—
Fu Ling	14.37	—
Ye Jiao Teng	15.48	—
Jiang Huang	15.70	—
Yin Yang Huo	16.12	—
Fo Shou	16.30	—
Ji Xue Teng	16.51	—
Gua Lou	16.97	—
Ma Huang Gen	17.48	—
Yu Jin	17.81	—
Ren Dong Teng	18.83	—
Chen Pi	18.93	—
Chuan Xiong	19.82	—
Tian Men Dong	19.85	—
Tu Fu Ling	19.94	—
Pi Pa Ye	20.31	—
Yuan Zhi	20.56	—
Shan Yao	20.82	—
Ma Bian Cao	21.19	—
Lian Qiao	21.26	—
Yi Mu Cao	21.33	—
Wu Yao	22.23	—
Wang Bu Liu Xing	22.30	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xian He Cao	22.36	—
(Bai) Dou Kou	22.56	—
Gu Sui Bu	22.85	—
Sha Ren	23.10	—
Long Yan Rou	23.42	—
Dan Shen	24.41	—
Da Zao	24.68	—
Yan Hu Suo	24.78	—
Fu Shen	24.93	—
Xu Duan	25.20	—
Dan Dou Chi	25.45	—
Fu Zi	25.66	—
Mao Dong Qing	25.79	—
Dang Shen	26.51	—
Sang Bai Pi	26.57	—
Nü Zhen Zi	28.46	—
Qiang Huo	29.01	—
Yin Chen Hao	29.03	—
Jiao Gu Lan	29.37	—
San Qi	29.50	—
Huang Bo	29.59	—
Yu Zhu	30.03	—
Dang Gui Wei	30.23	—
Ma Huang	30.33	—
Mang Xiao	30.38	—
Zhi Ke	30.99	—
Bu Gu Zhi	31.32	—
E Zhu	31.50	—
Zhi Mu	31.52	—
Shan Yu Rou	31.83	—
Chai Hu	31.91	—
Dang Gui	32.12	—
Hong Jing Tian	32.52	—
Fang Ji	32.74	—
Cang Zhu	33.51	—
Jin Yin Hua	33.54	—
Xiang Fu	33.59	—
Wu Wei Zi	34.16	—
Hou Po	34.24	—
Che Qian Zi	34.38	—
Sang Ji Shend	34.64	—
He Shou Wu	34.73	—
Fang Feng	35.16	—
Huang Jing	35.93	—
Gou Qi Zi	36.46	—
Zi Su Zi	36.92	—
Lu Gen	37.73	—
Huang Qin	37.76	—
Wu Zhu Yu	38.25	—
Hu Zhang	38.37	—
Cang Er Zi	38.37	—
Huang Lian	39.26	—
She Gan	39.96	—
Suan Zao Ren	40.18	—
Niu Bang Zi	40.37	—
Mu Dan Pi	40.86	—
Xia Ku Cao	40.96	—
Sang Zhi	41.28	—
Dan Zhu Ye	42.42	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Jiang	43.95	—
Ju Hua	44.07	—
Xin Yi	44.20	—
Qing Pi	44.22	—
Ge Gen	46.46	—
Ze Xie	46.66	—
(Shi) Chang Pu	47.64	—
Bai Zhu	47.68	—
Jie Geng	48.11	—
Du Zhong	48.48	—
Xi Xian Cao	49.66	—
Dong Gua Zi	50.43	—

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	14.92
62646	62646	0.00	3.40
62403	62403	0.00	15.25
62645	62645	0.00	3.64

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 2.1-2020-01

Validated substance/substance group **He Shou Wu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60101-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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	G198H0718421	3
PhytoComm	He Shou Wu	G198H0718222	1
PhytoComm	He Shou Wu	G198H0718421	1

- 846 spectra from 12 *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 *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	25 641
Type B	0	80	0	13 087
Type C	2	0	7	844

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.7674 % (> 99.1791 %)	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
Xian He Cao	10.58	—
Gu Sui Bu	13.74	—
Dan Zhu Ye	13.88	—
Sang Ji Shend	14.56	—
Xu Duan	14.59	—
Wu Yao	14.61	—
Yu Jin	15.62	—
He Huan Pi	15.99	—
Jing Jie	16.67	—
Sang Bai Pi	17.20	—
Ma Bian Cao	17.31	—
Hong Jing Tian	17.76	—
Lu Gen	18.89	—
Chen Pi	20.18	—
Nü Zhen Zi	21.79	—
Fang Feng	23.10	—
Hou Po	23.24	—
Yi Mu Cao	23.38	—
Bu Gu Zhi	23.58	—
Yin Chen Hao	23.94	—
Yin Yang Huo	23.95	—
Gui Zhi	24.69	—
Sha Ren	24.77	—
Ji Li	26.24	—
Dan Shen	26.30	—
Ge Gen	27.00	—
Huang Jing	28.50	—
Mu Dan Pi	29.15	—
Wu Zhu Yu	29.47	—
Lian Qiao	29.47	—
Jin Qian Cao	29.78	—
Fu Zi	29.87	—
Jiang Huang	29.91	—
Mang Xiao	29.92	—
Qiang Huo	30.19	—
Gou Teng	31.03	—
Du Zhong	33.40	—
Chuan Xiong	33.88	—
Huang Bo	34.46	—
Gou Qi Zi	35.22	—
Zhi Shi	35.86	—
She Gan	36.45	—
Xiang Fu	37.39	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
E Zhu	38.07	—
(Shi) Chang Pu	38.14	—
Ju Hua	39.06	—
(Bai) Dou Kou	39.31	—
Gan Jiang	39.47	—
Bai Mao Gen	39.48	—
Chai Hu	40.37	—
Sang Zhi	40.43	—
Pu Gong Ying	41.07	—
Du Huo	41.07	—
(Sheng) Di Huang	41.20	—
San Leng	41.46	—
Zi Wan	41.54	—
Huang Lian	42.01	—
Hu Zhang	42.02	—
Ba Ji Tian	42.55	—
Tu Fu Ling	42.72	—
Dang Gui	43.09	—
Wu Jia Pi	44.18	—
Huang Qi	44.81	—
Shan Yao	44.87	—
Qing Pi	45.16	—
Gua Lou	46.09	—
Huo Ma Ren	46.42	—
Sang Ye	46.46	—
Zi Su Zi	48.40	—
Xia Ku Cao	48.77	—
Ku Shen	49.02	—
Hua Shi	49.59	—
Bai Hua She She Cao	49.73	—
Dang Gui Wei	49.97	—

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
62745	62745	0.00	10.58
62301	62301	0.00	13.88
62746	62746	0.00	10.84
62302	62302	0.00	13.74

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 2.1-2020-01

Validated substance/substance group **Hong Hua**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60206-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Hong Hua	1	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

Validation samples

A total of 39 821 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 *Hong Hua*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	G057H0901221	1
Phytocomm	Hong Hua	g057h0901121	1

- 851 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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	10.54	—
Shan Yu Rou	12.45	—
Wu Wei Zi	13.54	—
Bai Zhu	15.79	—
Jiao Gu Lan	15.83	—
Zi Hua Di Ding	16.73	—
Jin Yin Hua	17.57	—
Fang Ji	17.65	—
Qing Pi	17.82	—
Zhi Ke	18.34	—
Sang Ye	18.36	—
Mao Dong Qing	22.17	—
Suan Zao Ren	22.64	—
Chai Hu	22.75	—
Cang Zhu	22.80	—
Dan Shen	23.04	—
Gua Lou	23.93	—
Wu Mei	24.22	—
Mu Dan Pi	24.53	—
Hong Jing Tian	25.06	—
Chuan Xiong	25.13	—
Huang Bai	25.47	—
Bai Bian Dou	26.48	—
Du Zhong	26.51	—
Ji Li	27.14	—
Lian Qiao	27.19	—
Dang Shen	27.32	—
Chuang Mu Xiang	27.39	—
Chen Pi	27.40	—
Ren Dong Teng	27.76	—
Cang Er Zi	28.04	—
Gan Cao	29.75	—
Gui Zhi	29.90	—
Qiang Huo	30.01	—
Gou Qi Zi	30.45	—
Yuan Zhi	30.63	—
Yi Mu Cao	30.79	—
Pi Pa Ye	31.50	—
Yan Hu Suo	31.82	—
Ze Xie	32.31	—
Hou Po	33.13	—
Bai Shao Yao	33.38	—
Dan Dou Chi	33.85	—
Zhe Bei Mu	33.91	—
Chi Shao (Yao)	34.07	—
Ye Jiao Teng	34.13	—
Ma Huang	34.31	—
Yin Yang Huo	34.74	—
Fu Zi	35.76	—
Di Gu Pi	35.84	—
Pu Gong Ying	35.97	—
Sang Zhi	36.21	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Lian	36.55	—
Lian Zi	36.84	—
He Huan Pi	36.95	—
Xi Xian Cao	37.03	—
Huang Qin	37.10	—
Zi Wan	37.64	—
San Qi	37.80	—
Bai Zi Ren	37.83	—
Yu Zhu	37.88	—
Shen Qu	38.13	—
Bai Xian Pi	38.15	—
Dang Gui	38.17	—
Tu Fu Ling	38.77	—
Long Yan Rou	39.05	—
Xia Ku Cao	39.21	—
(Shi) Chang Pu	39.69	—
Ling Zhi	39.99	—
Tai Zi Shen	40.25	—
Zhi Mu	40.25	—
Bai Mao Gen	40.48	—
San Leng	40.55	—
Huo Ma Ren	41.65	—
Dang Gui Wei	41.72	—
Niu Bang Zi	42.35	—
Mang Xiao	42.49	—
Tao Ren	42.93	—
Ji Xue Teng	43.29	—
Che Qian Zi	43.43	—
Ku Shen	43.61	—
Gou Teng	44.52	—
Jiang Huang	45.39	—
Ban Xia (Jiang)	47.00	—
Sheng Jiang	47.19	—
Yu Jin	47.96	—
Jie Geng	48.98	—
Xin Yi	48.99	—
Wu Yao	49.00	—
Fu Ling	49.34	—
Tian Hua Fen	49.43	—
Da Zao	49.76	—
Gan Jiang	50.16	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/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
62664	62664	0.00	10.54
62663	62663	0.00	10.70

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 2.1-2020-01

Validated substance/substance group **Hong Jing Tian**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60031-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Hong Jing Tian	3	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	G237H0906621	61911	40	from supplier
PhytoComm	Hong Jing Tian	G237H0906621	61912	40	from supplier
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 39 821 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 *Hong Jing Tian*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G237H0906621	61911 [†]	20
PhytoComm	Hong Jing Tian	G237H0906621	61912 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	G237H0906321	1
Phytocomm	Hong Jing Tian	G237H0906221	1

- 851 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 *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	240	0	25 561
Type B	0	118	2	13 047
Type C	0	0	2	851

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	98.3333 % (> 95.8333 %)
Type C	100.0000 % (> 98.8279 %)	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 *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	4.63	—
Mu Dan Pi	8.22	—
Dan Shen	9.32	—
Chuan Xiong	9.81	—
Hou Po	10.62	—
Zi Hua Di Ding	11.14	—
Gua Lou	12.24	—
Pi Pa Ye	12.87	—
Gou Qi Zi	13.24	—
Yuan Zhi	13.76	—
Ren Dong Teng	14.14	—
Fu Zi	14.44	—
Cang Zhu	15.45	—
Yin Yang Huo	15.58	—
Jiao Gu Lan	15.60	—
Bai Shao Yao	15.67	—
Ma Huang	15.67	—
Zhi Ke	16.46	—
Shan Yao	16.73	—
Ye Jiao Teng	16.87	—
Sang Ji Shend	17.29	—
Lian Qiao	17.35	—
Bai Bian Dou	17.44	—
Chai Hu	17.61	—
Yan Hu Suo	18.09	—
Qiang Huo	18.17	—
Wu Wei Zi	18.19	—
Fang Ji	18.88	—
Chen Pi	19.34	—
Bai Zhu	19.62	—
Qing Pi	19.90	—
Bai Xian Pi	20.19	—
Lian Zi	20.19	—
He Shou Wu	20.23	—
Shan Yu Rou	20.25	—
Xu Duan	20.48	—
Shen Qu	20.75	—
Dan Dou Chi	20.79	—
Ji Li	20.85	—
Mao Dong Qing	21.04	—
Gu Sui Bu	21.38	—
Tu Fu Ling	21.96	—
Chuang Mu Xiang	22.03	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhe Bei Mu	22.26	—
He Huan Pi	22.35	—
Ji Xue Teng	22.41	—
Di Gu Pi	23.04	—
Nü Zhen Zi	24.32	—
Gan Cao	24.47	—
San Qi	24.54	—
Wu Yao	24.60	—
Tao Ren	24.73	—
Long Yan Rou	25.22	—
Yin Chen Hao	25.30	—
Ling Zhi	25.48	—
Gou Teng	25.82	—
Huang Bai	25.91	—
Dang Gui	26.00	—
Tai Zi Shen	26.00	—
Dang Gui Wei	26.01	—
Che Qian Zi	26.12	—
Lu Gen	26.61	—
Yu Zhu	26.78	—
Tian Hua Fen	26.81	—
Huang Lian	27.21	—
Da Zao	27.43	—
Huang Qin	27.51	—
Yu Jin	27.62	—
Zhi Mu	27.64	—
Fang Feng	27.71	—
Ge Gen	27.74	—
Xian He Cao	27.92	—
Dang Shen	28.01	—
Gui Zhi	28.69	—
Rou Gui	28.78	—
Sang Bai Pi	28.92	—
Huo Ma Ren	29.12	—
Du Zhong	29.28	—
Huang Bo	29.41	—
Yi Mu Cao	29.44	—
Bai Zi Ren	30.12	—
Ju Hua	30.47	—
Sheng Jiang	30.65	—
Suan Zao Ren	31.39	—
Cang Er Zi	31.73	—
Ci Wu Jia	32.05	—
Chi Shao (Yao)	32.18	—
Sha Ren	32.48	—
Sang Zhi	32.74	—
Zhi Shi	33.01	—
Ze Xie	33.47	—
Fu Xiao Mai	33.51	—
Ban Xia (Jiang)	34.42	—
Fu Ling	34.87	—
Ma Bian Cao	34.95	—
Wu Zhu Yu	35.50	—
Jie Geng	36.60	—
Sang Ye	36.75	—
Wei Ling Xian	37.07	—
Mang Xiao	37.42	—
San Leng	37.46	—
Ma Huang Gen	38.04	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Hong Hua	38.14	—
Huang Jing	38.24	—
Dan Zhu Ye	38.35	—
Pu Gong Ying	38.63	—
Jing Jie	39.39	—
Hu Zhang	39.45	—
Bai Mao Gen	39.91	—
Yi Yi Ren	40.20	—
Xiang Fu	40.69	—
She Gan	41.21	—
Zhu Ru	41.68	—
Bu Gu Zhi	41.83	—
Fo Shou	42.77	—
(Shi) Chang Pu	43.48	—
Fu Shen	44.04	—
Chuan Lian Zi	44.65	—
Jiang Huang	45.63	—
Ban Zhi Lian	45.78	—
Zi Wan	46.10	—
Tian Men Dong	47.20	—
Du Huo	47.68	—
Xia Ku Cao	48.84	—
Ku Shen	48.86	—
Huang Qi	49.67	—

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
62639	62639	0.00	4.92
62305	62305	0.00	15.62
61912	61912	0.00	12.12
62640	62640	0.00	4.63
62306	62306	0.00	15.49
61911	61911	0.00	11.76

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 2.1-2020-01

Validated substance/substance group **Hou Po**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50289-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Hou Po	2	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	G154H0917822	62353	40	from supplier
PhytoComm	Hou Po	G154H0917822	62354	40	from supplier
PhytoComm	Hou Po	G154HS184SH1	62613	40	from supplier
PhytoComm	Hou Po	G154HS184SH1	62614	40	from supplier

Validation samples

A total of 39 821 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 *Hou Po*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G154H0917822	62353 [†]	20
PhytoComm	Hou Po	G154H0917822	62354 [†]	20
PhytoComm	Hou Po	G154HS184SH1	62613 [†]	20
PhytoComm	Hou Po	G154HS184SH1	62614 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	1	160	0	25 640
Type B	0	80	0	13 087
Type C	1	0	1	851

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	99.9927 % (> 99.9778 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.8837 % (> 99.3009 %)	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 *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	6.85	—
Yu Jin	7.34	—
Dan Dou Chi	7.34	—
Yan Hu Suo	7.45	—
Ren Dong Teng	7.76	—
Shen Qu	8.02	—
Mao Dong Qing	8.34	—
Shan Yao	8.49	—
Yin Yang Huo	8.62	—
Ling Zhi	8.88	—
Bai Bian Dou	8.91	—
Yuan Zhi	9.58	—
Bai Xian Pi	9.89	—
Pi Pa Ye	9.91	—
Che Qian Zi	10.10	—
Chuan Xiong	10.55	—
Qiang Huo	10.99	—
Ji Li	12.29	—
Dan Shen	12.48	—
Fang Ji	12.81	—
Tu Fu Ling	13.26	—
Lian Zi	14.28	—
Hong Jing Tian	14.33	—
Qing Pi	15.21	—
Lian Qiao	15.66	—
Jiao Gu Lan	15.92	—
Du Zhong	16.49	—
Wu Yao	16.51	—
Mu Dan Pi	16.59	—
Fu Zi	16.71	—
Ye Jiao Teng	16.93	—
Jin Yin Hua	16.98	—
Zi Hua Di Ding	18.06	—
Ma Huang	18.35	—
Ma Huang Gen	18.64	—
Ban Xia (Jiang)	18.77	—
Fang Feng	18.86	—
Bai Shao Yao	18.91	—
He Shou Wu	19.25	—
Ji Xue Teng	19.42	—
Gua Lou	19.70	—
Yi Mu Cao	19.80	—
He Huan Pi	20.16	—
Gou Teng	20.79	—
Gu Sui Bu	20.83	—
Ju Hua	20.85	—
Xian He Cao	21.14	—
Sheng Jiang	21.23	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tian Hua Fen	21.32	—
Fu Xiao Mai	21.34	—
Huo Ma Ren	21.73	—
Rou Gui	21.89	—
Chen Pi	22.09	—
Fu Ling	22.21	—
Wu Wei Zi	22.28	—
Cang Zhu	22.93	—
Gui Zhi	23.39	—
Yi Yi Ren	23.63	—
Huang Lian	23.71	—
Sha Ren	23.90	—
Bai Zi Ren	24.73	—
Huang Bo	25.15	—
Tao Ren	25.49	—
Wu Zhu Yu	25.89	—
Huang Bai	26.35	—
Gou Qi Zi	26.58	—
Tai Zi Shen	26.70	—
Lu Gen	26.75	—
Zhe Bei Mu	26.79	—
Di Gu Pi	27.25	—
Ma Bian Cao	27.56	—
Nü Zhen Zi	28.28	—
Da Zao	28.50	—
Sang Ji Shend	28.85	—
Zhi Ke	29.55	—
Xiang Fu	29.63	—
Suan Zao Ren	29.65	—
Xu Duan	30.09	—
Yin Chen Hao	30.42	—
Ci Wu Jia	31.73	—
(Shi) Chang Pu	31.85	—
San Qi	31.97	—
Zhu Ru	32.52	—
She Gan	32.96	—
Jiang Huang	33.10	—
Wei Ling Xian	33.36	—
Ge Gen	33.99	—
Shan Yu Rou	34.67	—
Yu Zhu	35.20	—
Mang Xiao	35.29	—
Sang Bai Pi	35.96	—
Huang Jing	36.14	—
Pu Gong Ying	36.26	—
Dang Gui	36.43	—
Du Huo	36.54	—
Long Yan Rou	37.01	—
Chuang Mu Xiang	37.52	—
Fu Shen	37.71	—
Ku Shen	37.74	—
Wang Bu Liu Xing	38.04	—
Sang Zhi	38.08	—
Huang Qin	38.26	—
Xie Bai	38.52	—
Dang Gui Wei	38.75	—
Fo Shou	38.75	—
Ze Xie	38.85	—
Zi Wan	39.55	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Zhu	39.73	—
Dang Shen	39.80	—
Jie Geng	40.12	—
Chi Shao (Yao)	40.80	—
Xi Xian Cao	40.81	—
Zhi Mu	41.64	—
Sang Ye	41.74	—
Tian Men Dong	41.78	—
Zhi Shi	41.99	—
Cang Er Zi	42.31	—
Hu Zhang	43.11	—
Dan Zhu Ye	43.23	—
Bai Mao Gen	44.14	—
Gan Cao	44.64	—
(Bai) Dou Kou	45.23	—
Ban Zhi Lian	46.35	—
Xia Ku Cao	46.35	—
Bu Gu Zhi	46.69	—
Xin Yi	46.75	—
Huang Qi	46.89	—
San Leng	47.74	—
Niu Bang Zi	48.37	—
Gan Jiang	48.62	—
Zi Su Zi	48.89	—
Hong Hua	49.38	—
E Zhu	49.47	—
Bai He	49.47	—

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	7.34
62614	62614	0.00	6.85
62353	62353	0.00	16.44
62354	62354	0.00	14.33

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 2.1-2020-01

Validated substance/substance group **Hu Zhang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60130-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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
Han Lian Cao	12.42	—
Xia Ku Cao	15.13	—
Sang Ye	15.43	—
Bai Hua She She Cao	17.37	—
Gui Zhi	20.90	—
Huang Lian	21.37	—
Pu Gong Ying	25.18	—
Sha Ren	25.85	—
Jiang Huang	26.86	—
Zhi Shi	27.00	—
Xi Xian Cao	27.11	—
Xin Yi	28.20	—
Yin Yang Huo	28.38	—
Fu Zi	29.37	—
Yi Mu Cao	31.02	—
Shan Yao	32.17	—
Nü Zhen Zi	33.00	—
Tu Fu Ling	33.87	—
Bu Gu Zhi	34.60	—
Sang Ji Shend	35.75	—
Rou Cong Rong	35.97	—
He Huan Pi	36.00	—
Yin Chen Hao	36.97	—
Huang Bo	37.05	—
Cang Er Zi	38.52	—
Gu Sui Bu	39.26	—
Qiang Huo	39.50	—
Gou Teng	39.72	—
Dan Zhu Ye	41.16	—
Wei Ling Xian	41.31	—
Shen Qu	42.68	—
Qing Pi	45.36	—
Dan Shen	45.80	—
Xu Duan	46.47	—
Mang Xiao	46.56	—
Wu Yao	46.97	—
Hong Jing Tian	47.42	—
Xuan Shen	48.47	—
Wu Zhu Yu	49.10	—
Ma Huang	49.94	—

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
62566	62566	0.00	12.64
62565	62565	0.00	12.42

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 2.1-2020-01

Validated substance/substance group **Hua Shi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50308-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	69	91	25 641
Type B	0	28	52	13 087
Type C	0	0	1	852

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 % (> 99.9701 %)	43.1250 % (> 41.2500 %)
Type B	100.0000 % (> 99.9422 %)	35.0000 % (> 31.2500 %)
Type C	100.0000 % (> 98.8343 %)	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	458.16	—

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	458.16
62637	62637	0.00	573.39
62426	62426	0.00	472.35
62638	62638	0.00	574.78

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 2.1-2020-01

Validated substance/substance group **Huang Bai**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60949-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huang Bai	1	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

Validation samples

A total of 39 821 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 *Huang Bai*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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 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.

[†]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
Qing Pi	7.78	—
Zhi Ke	8.70	—
Jiao Gu Lan	9.18	—
Dan Shen	10.56	—
Fang Ji	10.76	—
Qiang Huo	12.42	—
Zi Hua Di Ding	13.81	—
Hong Jing Tian	14.18	—
Du Zhong	15.33	—
Shan Yao	15.46	—
Chai Hu	16.14	—
Yin Yang Huo	16.78	—
Huang Lian	16.99	—
Wu Wei Zi	17.19	—
Jin Yin Hua	17.32	—
Sang Ye	17.50	—
Hou Po	17.68	—
Bai Xian Pi	20.73	—
Mao Dong Qing	20.83	—
Pi Pa Ye	20.84	—
Ye Jiao Teng	21.17	—
Yan Hu Suo	21.76	—
Mu Dan Pi	22.02	—
Ma Huang	22.59	—
Chuan Xiong	22.88	—
Bai Shao Yao	23.08	—
Bai Bian Dou	24.34	—
Shan Yu Rou	24.61	—
Ren Dong Teng	24.72	—
Yu Jin	24.72	—
Pu Gong Ying	25.39	—
Bai Zhu	25.63	—
Gui Zhi	25.91	—
Ji Xue Teng	25.92	—
Dan Dou Chi	26.26	—
Hong Hua	26.89	—
Fu Zi	27.84	—
Gua Lou	28.02	—
Lian Qiao	28.07	—
Di Gu Pi	28.16	—
Wu Mei	28.86	—
Lian Zi	29.03	—
Cang Er Zi	29.14	—
Ling Zhi	29.53	—
Cang Zhu	29.88	—
Ji Li	29.94	—
Xi Xian Cao	29.95	—
Tu Fu Ling	30.13	—
Gou Qi Zi	30.42	—
He Huan Pi	31.46	—
Xia Ku Cao	32.55	—
Che Qian Zi	33.04	—
Yuan Zhi	34.46	—
Huang Qin	34.74	—
Shen Qu	35.48	—
Yi Mu Cao	36.74	—
Tao Ren	37.48	—
Suan Zao Ren	38.57	—
Bai Mao Gen	38.68	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chi Shao (Yao)	39.20	—
Dang Gui	39.57	—
Rou Gui	39.77	—
Mang Xiao	40.36	—
Gan Cao	40.53	—
Jiang Huang	40.70	—
Zhe Bei Mu	40.73	—
Gou Teng	41.45	—
Sheng Jiang	42.45	—
Ban Zhi Lian	42.51	—
Ze Xie	42.62	—
Tian Hua Fen	42.69	—
San Qi	43.02	—
Chuang Mu Xiang	43.11	—
Chen Pi	43.14	—
Dang Shen	43.19	—
Nü Zhen Zi	44.36	—
Tai Zi Shen	44.45	—
Wu Yao	46.18	—
Bai Zi Ren	46.26	—
Xin Yi	46.26	—
Huo Ma Ren	46.41	—
Yu Zhu	46.48	—
Ban Xia (Jiang)	47.08	—
Fu Xiao Mai	47.49	—
Zhi Mu	47.98	—
Da Zao	48.09	—
Jie Geng	48.68	—
Ma Huang Gen	49.38	—
Dang Gui Wei	49.85	—
Sang Zhi	50.04	—

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	8.70
62698	62698	0.00	7.78

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 2.1-2020-01

Validated substance/substance group **Huang Bo**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60057-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Huang Bo; Phellodendri cortex chinensis

Special notes

When selecting the *Huang Bo* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huang Bo	1	0	3

Second-stage model

For differentiation of the substance/substance group *Huang Bo* 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 Bo*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Bo	G188H1203823	62593	40	from supplier
PhytoComm	Huang Bo	G188H1203823	62594	40	from supplier

Validation samples

A total of 39 821 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 *Huang Bo*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Huang Bo*.
- 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 Bo	G188H1203823	62593 [†]	20
PhytoComm	Huang Bo	G188H1203823	62594 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Huang Bo*.
- 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 Bo	h1203022	1
Phytocomm	Huang Bo	g188h1203123	1
Phytocomm	Huang Bo	G188H1203321	3
PhytoComm	Huang Bo	G188H1203321	1

- 847 spectra from 12 *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 Bo* 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 Bo* 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	25 721
Type B	0	40	0	13 127
Type C	2	1	5	845

The substance/substance group *Huang Bo* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.7010 % (> 99.1128 %)	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 Bo* 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	18.45	—
Yi Mu Cao	19.08	—
Huang Lian	31.44	—
Dan Zhu Ye	31.50	—
Qiang Huo	31.79	—
Wu Zhu Yu	33.03	—
Xu Duan	33.47	—
Gui Zhi	34.49	—
Wei Ling Xian	34.54	—
Yin Yang Huo	36.12	—
Gou Teng	37.61	—
Fu Zi	37.63	—
Bu Gu Zhi	39.45	—
Rou Cong Rong	40.31	—
Ma Bian Cao	40.46	—
Sha Ren	41.70	—
Xian He Cao	41.79	—
Hu Zhang	42.55	—
Wu Yao	43.15	—
Han Lian Cao	43.29	—
Hou Po	43.86	—
Ma Huang	44.57	—
Dan Shen	45.01	—
Yin Chen Hao	45.63	—
Jing Jie	46.21	—
Shen Qu	47.84	—
Hong Jing Tian	47.97	—
Cang Er Zi	49.98	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Bo* 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
62594	62594	0.00	18.45
62593	62593	0.00	18.49

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 2.1-2020-01

Validated substance/substance group **Huang Jing**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60108-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Huang Jing; Polygonati rhizoma

Special notes

When selecting the *Huang Jing* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huang Jing	1	0	1

Second-stage model

For differentiation of the substance/substance group *Huang Jing* 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 Jing*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Huang Jing	G196H1206721	62303	40	from supplier
PhytoComm	Huang Jing	G196H1206721	62304	40	from supplier

Validation samples

A total of 39 821 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 *Huang Jing*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Huang Jing*.
- 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 Jing	G196H1206721	62303 [†]	20
PhytoComm	Huang Jing	G196H1206721	62304 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Huang Jing*.
- 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	Huang Jing	G196H1206121	1
Phytocomm	Huang Jing	g196h1206121	1

- 851 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 *Huang Jing* 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 Jing* 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	25 721
Type B	0	40	0	13 127
Type C	0	0	2	851

The substance/substance group *Huang Jing* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 Jing* 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
Fang Feng	15.98	—
Gui Zhi	17.08	—
Chen Pi	18.01	—
He Huan Pi	21.79	—
Xu Duan	24.01	—
Gou Qi Zi	24.62	—
Sang Ji Shend	25.15	—
Du Huo	25.41	—
Chuan Xiong	25.48	—
Mang Xiao	25.99	—
Yi Mu Cao	26.35	—
Lian Qiao	26.52	—
Wu Yao	26.59	—
Ju Hua	27.64	—
Mu Dan Pi	27.68	—
Gan Jiang	27.72	—
Sang Bai Pi	28.00	—
Hong Jing Tian	28.05	—
Lu Gen	28.88	—
Dang Gui Wei	29.45	—
(Shi) Chang Pu	30.41	—
Gu Sui Bu	30.48	—
Jiang Huang	30.73	—
Ba Ji Tian	31.64	—
Zi Wan	32.32	—
Yu Jin	32.56	—
Gua Lou	32.84	—
Wu Zhu Yu	32.84	—
Hou Po	33.05	—
Ma Bian Cao	33.40	—
She Gan	33.51	—
Qiang Huo	34.25	—
Ji Li	34.30	—
San Leng	34.31	—
Tu Fu Ling	34.34	—
Yin Chen Hao	34.49	—
Yin Yang Huo	34.98	—
Xian He Cao	35.25	—
Huang Qi	35.35	—
Sang Zhi	35.62	—
Long Dan (Cao)	36.25	—
Nü Zhen Zi	36.33	—
Dan Shen	36.68	—
Dang Gui	36.82	—
Bai Mao Gen	37.41	—
Chai Hu	37.81	—
Huang Bo	38.61	—
Sha Ren	38.63	—
E Zhu	38.71	—
Huo Ma Ren	39.05	—
He Shou Wu	40.64	—
Cang Zhu	40.87	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xie Bai	41.59	—
Zhi Mu	41.88	—
Fu Zi	42.39	—
Qing Pi	42.57	—
Bai Bian Dou	42.77	—
Bu Gu Zhi	43.07	—
Jie Geng	44.15	—
Zi Su Zi	45.81	—
Di Gu Pi	46.17	—
Ge Gen	46.85	—
Lai Fu Zi	47.66	—
Ku Shen	47.86	—
Da Zao	47.87	—
Shan Yao	49.51	—
Chuan Lian Zi	49.67	—
Zhi Ke	49.88	—
Sheng Jiang	49.88	—
Zhi Shi	50.13	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Huang Jing* 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
62303	62303	0.00	16.71
62304	62304	0.00	15.98

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 2.1-2020-01

Validated substance/substance group **Huang Lian**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60219-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huang Lian	4	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	G079H1205621	61776	40	from supplier
PhytoComm	Huang Lian	G079H1205621	61819	40	from supplier
PhytoComm	Huang Lian	G079H1205622	62091	40	from supplier
PhytoComm	Huang Lian	G079H1205622	62092	40	from supplier
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

Validation samples

A total of 39 821 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 *Huang Lian*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 4 different batches.
- 25 481 spectra from a total of 315 batches from further 172 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 *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	G079H1205621	61776 [†]	20
PhytoComm	Huang Lian	G079H1205621	61819 [†]	20
PhytoComm	Huang Lian	G079H1205622	62091 [†]	20
PhytoComm	Huang Lian	G079H1205622	62092 [†]	20

continued on the next page

continued from previous page

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

- 13 007 spectra from a total of 316 batches from further 172 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	2
Phytocomm	Huang Lian	g079h1205321	1
Phytocomm	Huang Lian	G079H1205421	1
Phytocomm	Huang Lian	G079H1205521	2
PhytoComm	Huang Lian	G079H1205321	4
PhytoComm	Huang Lian	G079H1205521	2

- 841 spectra from 12 *Apo-Ident* customers from a total of 512 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	320	0	25 481
Type B	0	160	0	13 007
Type C	0	12	0	841

[†]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 *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 % (> 99.9701 %)	100.0000 % (> 98.1250 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 96.2500 %)
Type C	100.0000 % (> 98.8224 %)	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 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
Hu Zhang	13.51	—
Sang Ye	13.56	—
Pu Gong Ying	14.54	—
Xia Ku Cao	18.52	—
Shan Yao	22.13	—
Gui Zhi	22.14	—
Han Lian Cao	22.73	—
Ma Huang	24.07	—
Xin Yi	24.14	—
Bai Hua She She Cao	26.56	—
Zhi Ke	26.89	—
Huang Bai	27.08	—
Hou Po	28.06	—
Qiang Huo	28.09	—
Nü Zhen Zi	28.27	—
Qing Pi	28.91	—
Dan Shen	28.92	—
Yin Yang Huo	29.86	—
Tu Fu Ling	30.21	—
Zhi Shi	30.59	—
Xi Xian Cao	30.76	—
Wu Mei	30.82	—
Du Zhong	30.90	—
Ye Jiao Teng	31.00	—
Wu Wei Zi	31.13	—
Fu Zi	32.62	—
Yi Mu Cao	33.08	—
Jiao Gu Lan	33.30	—
Jiang Huang	33.53	—
Mao Dong Qing	34.16	—
Hong Jing Tian	34.39	—
Shen Qu	35.09	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sha Ren	35.67	—
Rou Cong Rong	36.03	—
Yu Jin	36.74	—
Bai Xian Pi	38.57	—
Pi Pa Ye	38.66	—
Cang Er Zi	38.72	—
Qing Hao	39.71	—
Gou Teng	39.80	—
Jin Yin Hua	39.89	—
Sang Ji Shend	40.51	—
Zi Hua Di Ding	40.64	—
Fang Ji	41.56	—
Chai Hu	42.30	—
Hong Hua	42.60	—
Wei Ling Xian	42.73	—
Bai Shao Yao	43.00	—
Bu Gu Zhi	43.25	—
Mang Xiao	43.82	—
Dan Zhu Ye	44.55	—
Huang Qin	44.60	—
Jing Jie	44.85	—
Shan Yu Rou	44.85	—
Bai Zhu	44.86	—
Ge Gen	44.91	—
Huang Bo	45.19	—
Wu Yao	45.34	—
Xu Duan	45.80	—
Che Qian Zi	46.16	—
Xuan Shen	47.58	—
He Huan Pi	47.66	—
Chuan Xiong	48.42	—
Gu Sui Bu	48.48	—
Yan Hu Suo	48.49	—
Di Gu Pi	48.50	—
Lian Qiao	48.94	—
Ze Xie	49.19	—
Bai Mao Gen	49.35	—
Xiang Fu	49.82	—
Chi Shao (Yao)	50.43	—

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
62707	62707	0.00	13.59
62091	62091	0.00	23.74

continued on the next page

continued from previous page

Sample ID	Reference sample ID	Distance to reference sample	Distance to next foreign sample
61776	61776	0.00	23.39
62287	62287	0.00	13.56
61819	61819	0.00	22.69
62288	62288	0.00	13.84
62708	62708	0.00	13.51
62092	62092	0.00	23.69

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 2.1-2020-01

Validated substance/substance group **Huang Qi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60154-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G040H1202721	62227	40	from supplier
PhytoComm	Huang Qi	G040H1202721	62228	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G040H1202721	62227 [†]	20
PhytoComm	Huang Qi	G040H1202721	62228 [†]	20
PhytoComm	Huang Qi	G040H1202822	62383 [†]	20
PhytoComm	Huang Qi	G040H1202822	62384 [†]	20
PhytoComm	Huang Qi	G040H1202823	62511 [†]	20
PhytoComm	Huang Qi	G040H1202823	62512 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	1
Phytocomm	Huang Qi	g040h1202321	2
Phytocomm	Huang Qi	G040H1202423	2
Phytocomm	Huang Qi	G040H1202521	2
PhytoComm	Huang Qi	G040H1202321	1
PhytoComm	Huang Qi	G040H1202621	1

- 843 spectra from 12 *Apo-Ident* customers from a total of 510 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	25 561
Type B	0	120	0	13 047
Type C	1	2	8	842

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	99.7674 % (> 99.1788 %)	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
Jie Geng	11.19	—
Bai Zhu	11.94	—
Gou Qi Zi	12.11	—
Mu Gua	14.40	—
She Gan	14.46	—
Chi Shao (Yao)	15.20	—
San Leng	16.00	—
Bai Shao Yao	17.03	—
Dang Gui Wei	17.18	—
Dang Shen	17.46	—
(Huai) Niu Xi	18.47	—
Sang Zhi	19.70	—
Di Gu Pi	19.99	—
Cang Zhu	21.25	—
Gui Zhi	22.94	—
Bai Mao Gen	23.27	—
Chuan Xiong	23.52	—
Yuan Zhi	24.19	—
Long Dan (Cao)	25.91	—
Chai Hu	26.01	—
Lu Gen	26.13	—
Mu Dan Pi	26.51	—
Ze Xie	28.09	—
Du Huo	28.28	—
Qin Jiao	28.76	—
Zi Wan	29.05	—
Zhi Mu	29.27	—
Bai He	30.54	—
Gua Lou	30.84	—
Xiang Fu	30.98	—
Dang Gui	32.19	—
(Shi) Chang Pu	32.71	—
Chen Pi	33.43	—
Mang Xiao	33.80	—
Lian Qiao	34.22	—
Long Yan Rou	34.26	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Cao	35.32	—
Zhi Ke	36.09	—
Gan Jiang	36.31	—
Mai Men Dong	36.40	—
Ban Zhi Lian	36.49	—
Tian Hua Fen	36.58	—
Shan Yu Rou	36.79	—
Tu Fu Ling	37.12	—
Da Zao	37.48	—
Qing Pi	37.64	—
San Qi	38.25	—
Sheng Jiang	38.77	—
Lai Fu Zi	39.12	—
Chuan Lian Zi	39.99	—
Ba Ji Tian	41.13	—
Lian Zi	41.16	—
Tian Men Dong	41.21	—
Sha Shen (Bei)	41.67	—
Zhe Bei Mu	41.97	—
Bai Xian Pi	42.05	—
Ji Li	42.23	—
Fang Feng	43.03	—
Huang Qin	43.63	—
Huang Lian	43.88	—
Wu Yao	44.12	—
Rou Gui	44.15	—
Jin Yin Hua	44.57	—
E Zhu	44.74	—
Shan Yao	44.82	—
Gou Teng	45.04	—
Dan Shen	45.32	—
Huo Ma Ren	46.65	—
Ji Xue Teng	46.72	—
Hong Jing Tian	47.21	—
Wu Wei Zi	47.83	—
Fu Zi	47.99	—
Yi Mu Cao	48.02	—
Wei Ling Xian	48.55	—
Ku Shen	48.78	—
Ye Jiao Teng	48.92	—
Dan Dou Chi	48.99	—
Fo Shou	49.74	—
Ju Hua	49.92	—

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
62512	62512	0.00	12.45
62384	62384	0.00	14.47
62383	62383	0.00	15.04
62228	62228	0.00	11.19
62227	62227	0.00	12.46
62511	62511	0.00	12.10

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 2.1-2020-01

Validated substance/substance group **Huang Qin**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60012-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huang Qin	2	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

Validation samples

A total of 39 821 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 Qin*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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

- 13 087 spectra from a total of 318 batches from further 172 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	G225H1204221	1
Phytocomm	Huang Qin	g225h1204121	1
Phytocomm	Huang Qin	G225H1204522	3
PhytoComm	Huang Qin	G225H1204221	1
PhytoComm	Huang Qin	G225H1204322	5

- 841 spectra from 12 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	12	841

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8224 %)	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 *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
Ma Huang	25.78	—
Dang Gui	26.65	—
Yu Zhu	28.09	—
Lian Qiao	29.43	—
Gui Zhi	30.40	—
Jiao Gu Lan	30.93	—
Zhi Ke	32.07	—
Gan Cao	32.21	—
Gou Qi Zi	33.00	—
Wu Wei Zi	33.28	—
Ye Jiao Teng	33.52	—
Hou Po	35.07	—
Tai Zi Shen	35.90	—
Bai Shao Yao	36.53	—
Yin Yang Huo	37.24	—
Fu Zi	37.80	—
Tian Hua Fen	38.46	—
Shan Yu Rou	38.58	—
Yan Hu Suo	38.74	—
Pi Pa Ye	39.00	—
Cang Zhu	39.57	—
Jin Yin Hua	39.98	—
Long Yan Rou	40.55	—
He Huan Pi	40.64	—
Zi Hua Di Ding	42.39	—
Chuang Mu Xiang	43.87	—
Chi Shao (Yao)	43.94	—
Dang Gui Wei	44.90	—
Huang Lian	45.13	—
Chen Pi	45.86	—
Zhi Mu	46.56	—
San Qi	46.61	—
Hong Jing Tian	47.28	—
Mang Xiao	47.46	—
Tao Ren	47.96	—
Dang Shen	47.99	—
Zhe Bei Mu	48.28	—
San Leng	48.49	—
Shan Yao	48.72	—
Hong Hua	48.97	—
Di Gu Pi	49.51	—
Ji Xue Teng	49.75	—

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
62282	62282	0.00	25.78
62709	62709	0.00	35.97
62281	62281	0.00	26.47
62710	62710	0.00	36.30

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 2.1-2020-01

Validated substance/substance group **Huo Ma Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60234-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Huo Ma Ren	2	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	G141H0449823	62411	40	from supplier
PhytoComm	Huo Ma Ren	G141H0449823	62412	40	from supplier
PhytoComm	Huo Ma Ren	G141HS061SP1	62711	40	from supplier
PhytoComm	Huo Ma Ren	G141HS061SP1	62712	40	from supplier

Validation samples

A total of 39 821 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 *Huo Ma Ren*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G141H0449823	62411 [†]	20
PhytoComm	Huo Ma Ren	G141H0449823	62412 [†]	20
PhytoComm	Huo Ma Ren	G141HS061SP1	62711 [†]	20
PhytoComm	Huo Ma Ren	G141HS061SP1	62712 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	0	160	0	25 641
Type B	0	79	1	13 087
Type C	0	0	2	851

The substance/substance group *Huo Ma 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	98.7500 % (> 95.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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	5.49	—
Bai Shao Yao	7.11	—
Gou Teng	7.22	—
Ban Xia (Jiang)	7.22	—
Bai Bian Dou	7.50	—
Tao Ren	7.91	—
Shen Qu	8.27	—
Gui Zhi	8.83	—
Lian Zi	8.86	—
Ling Zhi	8.95	—
Sheng Jiang	9.04	—
Rou Gui	9.12	—
Ci Wu Jia	9.46	—
Tai Zi Shen	9.72	—
Ji Li	9.77	—
Tian Hua Fen	9.83	—
Ji Xue Teng	10.08	—
Zhe Bei Mu	10.65	—
Ren Dong Teng	11.36	—
Yi Yi Ren	11.72	—
Bai Xian Pi	11.75	—
Fu Ling	11.93	—
Bai Zi Ren	12.32	—
Fu Xiao Mai	12.45	—
Ma Huang Gen	12.83	—
Ye Jiao Teng	13.44	—
Di Gu Pi	13.61	—
Zhu Ru	13.66	—
Cang Er Zi	13.83	—
Chuan Xiong	15.28	—
Chen Pi	15.48	—
Fo Shou	15.67	—
Xie Bai	15.79	—
Lian Qiao	15.98	—
Gan Jiang	16.03	—
Yu Jin	16.84	—
Shan Yao	16.85	—
Pi Pa Ye	16.88	—
Dan Dou Chi	17.16	—
Yan Hu Suo	17.27	—
Gua Lou	17.34	—
Dan Shen	17.79	—
Fu Zi	18.58	—
Tu Fu Ling	18.97	—
Zi Su Zi	19.48	—
Yin Yang Huo	20.53	—
Mao Dong Qing	20.56	—
Wei Ling Xian	20.99	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tian Men Dong	21.10	—
Ma Huang	21.63	—
Da Zao	21.83	—
She Gan	22.23	—
Yuan Zhi	22.87	—
Hong Jing Tian	23.23	—
Wang Bu Liu Xing	23.32	—
Chai Hu	23.40	—
Jiao Gu Lan	24.03	—
Ku Shen	24.17	—
Yi Mu Cao	24.70	—
Zhi Zi	25.37	—
Long Yan Rou	25.40	—
Hou Po	25.64	—
Fu Shen	25.67	—
Dong Gua Zi	26.09	—
Fang Ji	26.48	—
Che Qian Zi	26.49	—
Jiang Huang	26.77	—
Jin Yin Hua	26.93	—
Yu Zhu	27.13	—
Dang Gui Wei	28.19	—
Gou Qi Zi	28.37	—
Sang Zhi	28.98	—
Cang Zhu	29.13	—
San Qi	29.39	—
Wu Wei Zi	29.75	—
Mang Xiao	30.42	—
Zhi Ke	31.11	—
(Shi) Chang Pu	31.12	—
Xiang Fu	31.24	—
Dang Shen	31.93	—
Qiang Huo	31.96	—
Zhi Mu	32.14	—
Qing Pi	32.40	—
Shan Yu Rou	32.93	—
(Huai) Niu Xi	33.87	—
Jie Geng	34.08	—
Ze Xie	34.18	—
Dang Gui	34.49	—
Huang Qi	35.21	—
(Bai) Dou Kou	35.83	—
Huang Qin	35.93	—
Huang Lian	36.08	—
Niu Bang Zi	37.13	—
Nü Zhen Zi	37.22	—
Huang Bo	37.36	—
Mu Gua	37.55	—
Zi Wan	37.77	—
Suan Zao Ren	37.96	—
Sha Ren	38.13	—
Fang Feng	38.37	—
Du Zhong	39.21	—
Wu Mei	40.22	—
Mu Dan Pi	40.45	—
E Zhu	40.48	—
Du Huo	41.05	—
Chuan Lian Zi	42.12	—
Pu Gong Ying	42.36	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Bian Cao	43.43	—
Xu Duan	43.62	—
Bai Zhu	43.93	—
Bu Gu Zhi	44.26	—
Sha Shen (Bei)	44.53	—
Bai Mao Gen	44.70	—
Gan Cao	46.25	—
Chi Shao (Yao)	46.47	—
Zi Hua Di Ding	46.88	—
Lai Fu Zi	47.10	—
Wu Yao	48.80	—
Ba Ji Tian	49.19	—

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
62712	62712	0.00	5.49
62412	62412	0.00	14.19
62411	62411	0.00	13.83
62711	62711	0.00	5.94

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 2.1-2020-01

Validated substance/substance group **Ji Li**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60110-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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

- 13 047 spectra from a total of 317 batches from further 172 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

- 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 *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	0	240	0	25 561
Type B	2	120	0	13 045
Type C	2	0	1	850

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	100.0000 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	99.9855 % (> 99.9565 %)	100.0000 % (> 95.0000 %)
Type C	99.7674 % (> 99.1846 %)	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
Tian Hua Fen	5.18	—
Lian Qiao	5.62	—
Bai Shao Yao	6.65	—
Bai Bian Dou	7.01	—
Huo Ma Ren	7.33	—
Tai Zi Shen	7.46	—
Bai Zi Ren	7.80	—
Ling Zhi	7.85	—
He Huan Pi	8.07	—
Shen Qu	8.23	—
Rou Gui	8.54	—
Fu Zi	8.68	—
Lian Zi	8.81	—
Ban Xia (Jiang)	8.91	—
Zhe Bei Mu	8.98	—
Ci Wu Jia	9.31	—
Ren Dong Teng	9.45	—
Sang Zhi	9.46	—
Pi Pa Ye	10.03	—
Zhu Ru	10.25	—
Fu Xiao Mai	10.40	—
Gua Lou	10.74	—
Shan Yao	10.82	—
Fu Ling	11.01	—
Hou Po	11.60	—
Tao Ren	11.66	—
Sheng Jiang	11.92	—
Gui Zhi	11.99	—
Chuan Xiong	12.02	—
Di Gu Pi	12.91	—
Mao Dong Qing	13.12	—
Gou Teng	13.23	—
Jiang Huang	13.56	—
Chai Hu	13.61	—
Dan Shen	14.33	—
Tu Fu Ling	14.38	—
Yu Jin	14.44	—
Yi Yi Ren	14.63	—
E Zhu	14.92	—
Che Qian Zi	15.05	—
Chen Pi	15.47	—
Ye Jiao Teng	15.62	—
Jiao Gu Lan	15.65	—
Bai Xian Pi	16.05	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	16.06	—
Ji Xue Teng	16.34	—
Cang Zhu	17.14	—
Jin Yin Hua	17.41	—
Da Zao	17.53	—
Sha Ren	17.60	—
Dan Dou Chi	17.71	—
Ba Ji Tian	18.09	—
(Bai) Dou Kou	18.28	—
Yan Hu Suo	18.51	—
Dang Gui Wei	18.78	—
San Qi	18.81	—
Fo Shou	19.11	—
Yi Mu Cao	19.91	—
Gan Jiang	19.96	—
Yin Yang Huo	20.07	—
Chuan Lian Zi	20.13	—
Suan Zao Ren	20.22	—
Wang Bu Liu Xing	20.55	—
Ma Huang Gen	20.65	—
Wu Wei Zi	20.93	—
Tian Men Dong	21.24	—
Fang Ji	21.27	—
Long Dan (Cao)	22.19	—
Lu Gen	22.20	—
Wu Yao	22.50	—
Hong Jing Tian	22.72	—
Yu Zhu	22.76	—
Dang Shen	22.93	—
Sha Shen (Bei)	22.97	—
Du Huo	23.41	—
Mu Dan Pi	23.43	—
Qing Pi	23.51	—
Xian He Cao	23.53	—
Mang Xiao	23.61	—
Zi Su Zi	23.69	—
Qiang Huo	23.99	—
Ma Huang	24.36	—
She Gan	24.43	—
Shan Yu Rou	24.77	—
Dang Gui	25.55	—
Gou Qi Zi	26.10	—
Zhi Mu	26.40	—
Xiang Fu	26.82	—
Du Zhong	27.34	—
(Huai) Niu Xi	27.48	—
Long Yan Rou	27.94	—
Fang Feng	27.99	—
Cang Er Zi	28.39	—
Xie Bai	28.53	—
Niu Bang Zi	28.79	—
Zhi Ke	29.75	—
(Shi) Chang Pu	29.94	—
Huang Qin	30.16	—
Lai Fu Zi	31.14	—
Fu Shen	31.61	—
Huang Lian	32.95	—
San Leng	32.97	—
Chuang Mu Xiang	33.12	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Wan	33.64	—
Zi Hua Di Ding	33.99	—
Bai Zhu	36.00	—
Bai Mao Gen	36.35	—
Gan Cao	36.47	—
Bai He	37.30	—
Jie Geng	37.84	—
Zhu Ling	38.69	—
Huang Qi	38.76	—
Ze Xie	39.24	—
Xu Duan	39.61	—
Xin Yi	39.83	—
Ju Hua	40.44	—
Chi Shao (Yao)	40.91	—
Sang Bai Pi	41.03	—
Sang Ji Shend	41.19	—
He Shou Wu	42.60	—
Huang Jing	43.08	—
Huang Bai	44.55	—
Nü Zhen Zi	44.77	—
Huang Bo	44.90	—
Qin Jiao	45.12	—
Ma Bian Cao	45.69	—
Mai Ya	46.10	—
Mo Yao	47.01	—
Ku Shen	47.47	—
Mu Gua	48.26	—
Yin Chen Hao	48.74	—
Gu Sui Bu	48.97	—
Bu Gu Zhi	49.16	—
Dong Gua Zi	49.95	—
Sang Ye	50.45	—

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
62721	62721	0.00	5.18
62319	62319	0.00	9.90
62546	62546	0.00	6.19
62320	62320	0.00	9.46
62545	62545	0.00	5.62
62722	62722	0.00	5.94

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 2.1-2020-01

Validated substance/substance group **Ji Xue Teng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60099-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 849 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 *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	4	157	3	25 637
Type B	3	80	0	13 084
Type C	0	0	4	849

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	99.9903 % (> 99.9754 %)	98.1250 % (> 96.2500 %)
Type B	99.9855 % (> 99.9566 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8246 %)	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
Ye Jiao Teng	4.55	—
Rou Gui	6.05	—
Ma Huang Gen	6.35	—
Ren Dong Teng	8.52	—
Gou Teng	9.30	—
Fo Shou	10.85	—
Huo Ma Ren	11.13	—
Tai Zi Shen	11.24	—
Lian Zi	11.97	—
Bai Shao Yao	12.09	—
Bai Xian Pi	12.46	—
He Huan Pi	14.08	—
Gui Zhi	14.20	—
Chen Pi	15.22	—
Dan Dou Chi	15.54	—
Ling Zhi	15.61	—
Sheng Jiang	15.91	—
Shen Qu	15.95	—
Yi Yi Ren	16.07	—
Bai Bian Dou	16.62	—
Tao Ren	16.74	—
Fu Shen	16.80	—
Tu Fu Ling	16.85	—
Ban Xia (Jiang)	17.05	—
Dan Shen	17.25	—
Ci Wu Jia	17.79	—
Fu Ling	18.55	—
Ma Huang	18.66	—
Yan Hu Suo	18.66	—
Yin Yang Huo	20.07	—
Da Zao	20.67	—
Zhe Bei Mu	20.88	—
Di Gu Pi	21.12	—
Yu Jin	21.56	—
Tian Men Dong	21.84	—
Zhu Ru	22.14	—
Chuan Xiong	22.23	—
Tian Hua Fen	22.55	—
Ji Li	23.13	—
Lian Qiao	23.39	—
Long Yan Rou	25.04	—
Fu Xiao Mai	25.53	—
Bai Zi Ren	25.60	—
Hong Jing Tian	25.64	—
Mao Dong Qing	25.81	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Pi Pa Ye	25.91	—
Gua Lou	26.29	—
Yuan Zhi	26.57	—
Jin Yin Hua	26.73	—
Jiao Gu Lan	28.41	—
Shan Yao	29.01	—
Fu Zi	29.04	—
Chai Hu	29.47	—
Zhi Ke	30.48	—
Fang Ji	30.56	—
Yu Zhu	30.79	—
Hou Po	31.30	—
Cang Zhu	32.76	—
San Qi	33.66	—
Huang Lian	33.70	—
Qiang Huo	33.93	—
Zhi Mu	34.89	—
Dang Gui Wei	35.28	—
Wu Wei Zi	35.45	—
Mang Xiao	35.80	—
Wang Bu Liu Xing	36.41	—
Huang Qin	36.55	—
Dang Shen	37.32	—
Shan Yu Rou	38.16	—
Che Qian Zi	38.70	—
Sang Zhi	38.78	—
Mu Dan Pi	39.05	—
Jie Geng	39.14	—
Gou Qi Zi	40.37	—
Zi Hua Di Ding	41.47	—
Du Zhong	42.92	—
Dang Gui	43.34	—
Gan Cao	43.82	—
Bai He	43.89	—
Chi Shao (Yao)	44.25	—
Bai Zhu	45.67	—
Qing Pi	46.09	—
Ze Xie	47.08	—
Ban Zhi Lian	47.34	—
San Leng	48.70	—

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	4.55
62723	62723	0.00	6.34
62560	62560	0.00	4.66
62724	62724	0.00	6.05

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 2.1-2020-01

Validated substance/substance group **Jiang Huang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60181-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Jiang Huang	2	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	G085H1716721	62373	40	from supplier
PhytoComm	Jiang Huang	G085H1716721	62374	40	from supplier
PhytoComm	Jiang Huang	G085H1716821	62519	40	from supplier
PhytoComm	Jiang Huang	G085H1716821	62520	40	from supplier

Validation samples

A total of 39 821 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 *Jiang Huang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G085H1716721	62373 [†]	20
PhytoComm	Jiang Huang	G085H1716721	62374 [†]	20
PhytoComm	Jiang Huang	G085H1716821	62519 [†]	20
PhytoComm	Jiang Huang	G085H1716821	62520 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G085H1716422	3
PhytoComm	Jiang Huang	G085H1716321	1
PhytoComm	Jiang Huang	G085H1716422	1

- 848 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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	2	3	2	846

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.5349 % (> 98.9469 %)	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 *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
Gui Zhi	7.68	—
Sha Ren	14.02	—
Fu Zi	15.73	—
Yi Mu Cao	16.35	—
Ji Li	18.01	—
Pu Gong Ying	20.92	—
(Bai) Dou Kou	20.98	—
He Huan Pi	21.14	—
Ze Xie	21.93	—
Yan Hu Suo	22.55	—
Wu Yao	23.42	—
Shan Yao	23.59	—
E Zhu	24.12	—
Cang Er Zi	24.16	—
Yin Yang Huo	25.18	—
Xi Xian Cao	25.81	—
Xiang Fu	26.86	—
Wu Mei	26.87	—
Tu Fu Ling	26.94	—
Qiang Huo	26.95	—
Huang Lian	27.04	—
Chen Pi	27.06	—
Gua Lou	27.86	—
Lu Gen	28.13	—
Yu Jin	28.57	—
Chai Hu	28.71	—
Sang Ye	28.74	—
Qing Pi	28.90	—
Wang Bu Liu Xing	29.22	—
Bu Gu Zhi	29.52	—
Wu Wei Zi	30.64	—
Du Zhong	30.78	—
Xin Yi	30.96	—
Mang Xiao	31.21	—
Zi Wan	32.14	—
Sang Zhi	32.47	—
Hou Po	32.53	—
Hu Zhang	32.58	—
Xia Ku Cao	32.71	—
Wei Ling Xian	33.06	—
Gou Qi Zi	33.63	—
Niu Bang Zi	33.97	—
Di Gu Pi	34.35	—
Zhi Ke	34.41	—
She Gan	34.66	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zi Su Zi	34.66	—
Xu Duan	34.73	—
Dan Shen	35.09	—
Han Lian Cao	35.74	—
Xian He Cao	35.75	—
Fang Ji	36.39	—
Huo Ma Ren	36.52	—
(Shi) Chang Pu	36.52	—
Ma Bian Cao	36.88	—
Huang Bo	36.89	—
Tian Hua Fen	37.03	—
Jiao Gu Lan	37.39	—
Zi Hua Di Ding	38.15	—
Sha Shen (Bei)	38.57	—
Hong Jing Tian	38.75	—
Dan Zhu Ye	38.86	—
Nü Zhen Zi	38.98	—
Mu Dan Pi	39.32	—
Ku Shen	40.03	—
Xie Bai	40.05	—
Gan Jiang	40.25	—
Huang Bai	40.56	—
Sang Bai Pi	41.02	—
Mao Dong Qing	41.03	—
Sang Ji Shend	41.18	—
Ye Jiao Teng	41.37	—
Che Qian Zi	41.66	—
Lian Qiao	41.79	—
Zhi Mu	42.16	—
Yin Chen Hao	42.41	—
Gu Sui Bu	42.43	—
Dan Dou Chi	42.86	—
Dang Gui Wei	43.23	—
Bai Hua She She Cao	43.83	—
Zhi Zi	43.97	—
Ba Ji Tian	44.07	—
He Shou Wu	45.23	—
Ma Huang	45.24	—
Chuan Xiong	45.24	—
Du Huo	45.80	—
Fang Feng	45.95	—
Dang Gui	46.04	—
Cang Zhu	46.22	—
Jin Yin Hua	46.36	—
Chuan Lian Zi	48.70	—
Bai Bian Dou	48.90	—
Bai Mao Gen	49.95	—
Pi Pa Ye	50.01	—

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
62374	62374	0.00	15.99
62519	62519	0.00	7.89
62373	62373	0.00	15.73
62520	62520	0.00	7.68

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 2.1-2020-01

Validated substance/substance group **Jiao Gu Lan**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60332-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Jiao Gu Lan	2	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

Validation samples

A total of 39 821 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 *Jiao Gu Lan*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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

- 850 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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	3	850

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8257 %)	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 *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	9.38	—
Shan Yao	10.50	—
Suan Zao Ren	10.67	—
Cang Zhu	10.73	—
Dan Shen	11.07	—
Bai Zhu	11.72	—
Zi Hua Di Ding	11.77	—
Mu Dan Pi	12.05	—
Bai Bian Dou	12.06	—
Lian Qiao	12.18	—
Wu Wei Zi	12.89	—
Ji Li	13.13	—
Hong Jing Tian	13.46	—
Shan Yu Rou	13.67	—
Qing Pi	14.12	—
Chai Hu	14.62	—
Mao Dong Qing	14.66	—
Fang Ji	14.87	—
Chuan Xiong	15.64	—
Bai Zi Ren	17.53	—
Gou Qi Zi	17.65	—
Pi Pa Ye	17.78	—
Gua Lou	18.01	—
Chen Pi	18.95	—
Yuan Zhi	19.01	—
Zhi Ke	19.51	—
Dang Shen	19.82	—
Hou Po	20.49	—
Ren Dong Teng	20.80	—
Ling Zhi	20.81	—
Ban Xia (Jiang)	21.06	—
Bai Shao Yao	21.24	—
Lian Zi	21.74	—
Cang Er Zi	22.02	—
Qiang Huo	22.77	—
Yan Hu Suo	22.94	—
Du Zhong	23.03	—
Hong Hua	23.22	—
Yin Yang Huo	23.89	—
Fu Zi	24.05	—
Shen Qu	24.57	—
Dan Dou Chi	24.62	—
Huang Bai	25.04	—
Ma Huang	25.06	—
Ye Jiao Teng	25.09	—
Di Gu Pi	25.59	—
Bai Xian Pi	26.37	—
Chuang Mu Xiang	26.66	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
San Qi	27.09	—
Zhe Bei Mu	27.51	—
Sang Ye	27.63	—
Yu Zhu	28.27	—
Huo Ma Ren	28.99	—
Ji Xue Teng	29.30	—
Fu Ling	29.38	—
He Huan Pi	29.43	—
Sheng Jiang	29.54	—
Tai Zi Shen	30.18	—
Fu Xiao Mai	30.26	—
Dang Gui	31.36	—
Zhi Mu	31.77	—
Huang Lian	32.20	—
Chi Shao (Yao)	32.44	—
Yu Jin	32.57	—
Rou Gui	33.04	—
Huang Qin	33.16	—
Gui Zhi	33.23	—
Gou Teng	33.45	—
Gan Cao	34.19	—
Che Qian Zi	34.24	—
Tao Ren	35.03	—
Long Yan Rou	35.50	—
Dang Gui Wei	35.78	—
Tu Fu Ling	35.95	—
Tian Hua Fen	36.03	—
Yi Mu Cao	36.17	—
Zhu Ru	37.37	—
Wu Mei	37.59	—
Sang Zhi	37.98	—
Pu Gong Ying	39.14	—
Bai Mao Gen	39.39	—
Da Zao	39.39	—
Xi Xian Cao	39.43	—
Ze Xie	39.74	—
Mang Xiao	40.51	—
Ci Wu Jia	41.28	—
Yi Yi Ren	42.45	—
Jie Geng	44.99	—
Chuan Lian Zi	45.19	—
Xia Ku Cao	46.38	—
San Leng	46.80	—
Zi Wan	47.44	—
Ma Huang Gen	47.78	—
Zi Su Zi	47.87	—
Niu Bang Zi	48.25	—

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
62442	62442	0.00	10.78
62441	62441	0.00	9.38
62666	62666	0.00	10.78
62665	62665	0.00	10.50

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 2.1-2020-01

Validated substance/substance group **Jie Geng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60034-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G193H1015621	61931	40	from supplier
PhytoComm	Jie Geng	G193H1015621	61932	40	from supplier
PhytoComm	Jie Geng	G193H1015721	62247	40	from supplier
PhytoComm	Jie Geng	G193H1015721	62248	40	from supplier
PhytoComm	Jie Geng	G193H1015821	62365	40	from supplier
PhytoComm	Jie Geng	G193H1015821	62366	40	from supplier

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G193H1015621	61931 [†]	20
PhytoComm	Jie Geng	G193H1015621	61932 [†]	20
PhytoComm	Jie Geng	G193H1015721	62247 [†]	20
PhytoComm	Jie Geng	G193H1015721	62248 [†]	20
PhytoComm	Jie Geng	G193H1015821	62365 [†]	20
PhytoComm	Jie Geng	G193H1015821	62366 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 848 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 *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	3	236	4	25 558
Type B	4	112	8	13 043
Type C	4	4	1	844

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	99.9938 % (> 99.9788 %)	98.3333 % (> 97.0833 %)
Type B	99.9834 % (> 99.9544 %)	93.3333 % (> 90.8333 %)
Type C	99.6382 % (> 99.0502 %)	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
Bai Zhu	3.68	—
Di Gu Pi	8.43	—
Huang Qi	9.06	—
Sang Zhi	9.83	—
Gou Qi Zi	10.51	—
Mu Gua	10.70	—
Dang Shen	10.80	—
(Huai) Niu Xi	11.92	—
She Gan	13.41	—
Dang Gui Wei	14.43	—
San Leng	15.74	—
Cang Zhu	16.15	—
Bai Mao Gen	16.22	—
Yuan Zhi	17.07	—
Chai Hu	18.48	—
Gui Zhi	18.64	—
San Qi	19.88	—
Chuan Xiong	20.11	—
Bai Shao Yao	20.26	—
Chi Shao (Yao)	21.60	—
Mai Men Dong	21.64	—
Zi Wan	21.82	—
Tian Hua Fen	21.83	—
Zhi Mu	22.05	—
Chuan Lian Zi	22.71	—
Du Huo	23.00	—
Bai He	23.09	—
Qin Jiao	23.64	—
Lu Gen	23.75	—
Gua Lou	24.22	—
Sheng Jiang	24.66	—
Long Dan (Cao)	25.30	—
Qing Pi	25.85	—
(Shi) Chang Pu	25.86	—
Lai Fu Zi	26.29	—
Mu Dan Pi	26.32	—
Gan Jiang	26.55	—
Xiang Fu	27.06	—
Tu Fu Ling	28.39	—
Dang Gui	28.68	—
Bai Xian Pi	29.19	—
Ji Li	29.93	—
Lian Zi	29.97	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mang Xiao	30.21	—
Long Yan Rou	30.74	—
Fang Feng	31.39	—
Rou Gui	31.85	—
Ban Zhi Lian	31.96	—
Ze Xie	32.17	—
Tian Men Dong	32.58	—
Shan Yao	32.86	—
E Zhu	33.89	—
Lian Qiao	34.12	—
Ba Ji Tian	35.53	—
Da Zao	36.35	—
Zhe Bei Mu	36.53	—
Chen Pi	36.63	—
Jin Yin Hua	36.71	—
Cang Er Zi	36.72	—
Huo Ma Ren	37.07	—
Fu Ling	37.19	—
Ye Jiao Teng	38.00	—
Zi Su Zi	38.07	—
Gou Teng	38.81	—
Zhi Ke	39.25	—
Ban Xia (Jiang)	39.27	—
Fo Shou	39.54	—
Sha Shen (Bei)	39.56	—
Hong Jing Tian	40.03	—
Bai Bian Dou	40.06	—
Ku Shen	40.74	—
Wu Wei Zi	40.95	—
Xie Bai	41.13	—
Shan Yu Rou	41.24	—
Ji Xue Teng	41.25	—
Huang Lian	41.39	—
Niu Bang Zi	41.50	—
Ju Hua	41.69	—
Fu Zi	41.70	—
Ren Dong Teng	42.42	—
Mai Ya	42.55	—
Yi Mu Cao	42.57	—
Ma Huang Gen	42.96	—
Wu Yao	43.22	—
Jiang Huang	43.42	—
Yi Yi Ren	43.57	—
Dan Shen	44.46	—
Zhu Ru	44.80	—
Dan Dou Chi	44.93	—
Qiang Huo	44.94	—
Gan Cao	45.34	—
He Huan Pi	45.92	—
Shen Qu	46.71	—
Huang Qin	46.87	—
Hou Po	47.17	—
Yin Yang Huo	47.85	—
Yu Jin	48.01	—
Tao Ren	48.79	—
Mao Dong Qing	48.79	—
Wei Ling Xian	49.62	—
Yan Hu Suo	50.02	—

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
62247	62247	0.00	3.68
62248	62248	0.00	4.18
62365	62365	0.00	8.42
62366	62366	0.00	8.43
61931	61931	0.00	9.53
61932	61932	0.00	9.06

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 2.1-2020-01

Validated substance/substance group **Jin Qian Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50293-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	1	852

The substance/substance group *Jin Qian 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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
Jing Jie	18.12	—
Pu Gong Ying	18.37	—
Wu Jia Pi	22.17	—
Du Zhong	23.45	—
Sha Ren	27.00	—
Sang Ji Shend	28.13	—
Dan Zhu Ye	28.41	—
Gou Teng	28.53	—
Yin Yang Huo	30.56	—
Huang Lian	30.91	—
Yin Chen Hao	31.99	—
Ge Gen	32.90	—
Xuan Shen	33.01	—
Xian He Cao	33.62	—
(Sheng) Di Huang	34.13	—
Nü Zhen Zi	35.00	—
Dan Shen	35.19	—
Qing Hao	36.37	—
Hong Jing Tian	36.54	—
Yu Xing Cao	37.13	—
Xu Duan	37.68	—
Ye Jiao Teng	38.01	—
Sang Bai Pi	38.06	—
He Shou Wu	39.91	—
Bu Gu Zhi	43.40	—
Zhi Shi	43.81	—
Gu Sui Bu	46.15	—
Qiang Huo	46.25	—
Hua Shi	48.36	—
Huang Bo	48.58	—
Ma Bian Cao	49.87	—

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
62466	62466	0.00	18.12
62465	62465	0.00	18.25

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 2.1-2020-01

Validated substance/substance group **Jin Yin Hua**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60350-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Jin Yin Hua	1	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

Validation samples

A total of 39 821 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 Yin Hua*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	80	0	25 721
Type B	2	40	0	13 125
Type C	0	0	0	853

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	99.9903 % (> 99.9615 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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 *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.

[†]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
Jiao Gu Lan	8.75	—
Mu Dan Pi	10.77	—
Shan Yu Rou	11.00	—
Chuan Xiong	11.36	—
Pi Pa Ye	11.65	—
Lian Qiao	11.77	—
Fu Zi	11.79	—
Gua Lou	12.20	—
Cang Zhu	12.36	—
Yan Hu Suo	12.55	—
Hou Po	12.96	—
Gou Qi Zi	13.35	—
Fang Ji	13.72	—
Wu Wei Zi	14.07	—
Shan Yao	14.34	—
Ren Dong Teng	15.37	—
Bai Bian Dou	15.47	—
Zhi Ke	16.76	—
Yuan Zhi	17.09	—
Bai Shao Yao	17.29	—
Qing Pi	17.31	—
San Qi	17.33	—
Hong Jing Tian	17.61	—
Chen Pi	17.99	—
Dan Shen	18.06	—
Ji Li	18.44	—
Ma Huang	18.61	—
Zi Hua Di Ding	18.67	—
Bai Zhu	18.80	—
Zhe Bei Mu	18.90	—
Dang Gui	19.09	—
Chai Hu	19.64	—
Chuang Mu Xiang	20.44	—
Dang Shen	20.66	—
Bai Zi Ren	21.10	—
Tai Zi Shen	21.37	—
Suan Zao Ren	21.59	—
Ye Jiao Teng	21.59	—
He Huan Pi	21.80	—
Yin Yang Huo	21.81	—
Yu Zhu	22.48	—
Mao Dong Qing	23.47	—
Huo Ma Ren	23.78	—
Ji Xue Teng	24.14	—
Lian Zi	24.19	—
Tian Hua Fen	24.95	—
Qiang Huo	25.21	—
Shen Qu	25.37	—
Bai Xian Pi	25.67	—
Gan Cao	25.72	—
Tao Ren	25.83	—
Tu Fu Ling	25.98	—
Di Gu Pi	26.34	—
Dang Gui Wei	27.01	—
Huang Qin	27.34	—
Dan Dou Chi	27.66	—
Zhi Mu	28.68	—
Du Zhong	29.26	—
Ling Zhi	29.53	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Teng	29.77	—
Che Qian Zi	30.07	—
Sheng Jiang	30.54	—
Huang Lian	30.71	—
Ban Xia (Jiang)	30.96	—
Sang Ye	32.13	—
Cang Er Zi	32.47	—
Fu Xiao Mai	32.70	—
Yu Jin	32.73	—
Gui Zhi	32.83	—
Hong Hua	33.14	—
Fu Ling	33.77	—
Chi Shao (Yao)	33.87	—
Long Yan Rou	34.00	—
Ci Wu Jia	34.16	—
Da Zao	34.35	—
Rou Gui	36.18	—
Huang Bai	36.32	—
Sang Zhi	36.55	—
Yi Yi Ren	37.17	—
Zhu Ru	38.23	—
Ze Xie	39.59	—
Mang Xiao	40.74	—
Bai Mao Gen	43.06	—
Yi Mu Cao	43.28	—
San Leng	43.37	—
Jie Geng	45.94	—
Ma Huang Gen	46.73	—
Pu Gong Ying	47.16	—
(Shi) Chang Pu	47.36	—
Xi Xian Cao	48.24	—
Zi Wan	49.18	—
Chuan Lian Zi	49.24	—

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
62578	62578	0.00	8.75
62577	62577	0.00	8.91

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 2.1-2020-01

Validated substance/substance group **Jing Jie**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60088-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Jing Jie	2	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

Validation samples

A total of 39 821 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 *Jing Jie*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	0	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	2	851

The substance/substance group *Jing Jie* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Du Zhong	8.45	—
Dan Zhu Ye	8.48	—
Yin Chen Hao	10.35	—
Sang Bai Pi	14.16	—
Bu Gu Zhi	16.68	—
He Shou Wu	16.99	—
Sha Ren	18.29	—
Jin Qian Cao	18.37	—
Xu Duan	18.39	—
Wu Jia Pi	18.84	—
Sang Ji Shend	18.98	—
Nü Zhen Zi	21.89	—
Ma Bian Cao	21.92	—
Xian He Cao	23.37	—
(Sheng) Di Huang	26.75	—
Gu Sui Bu	26.83	—
Yin Yang Huo	27.62	—
Pu Gong Ying	28.84	—
Yi Mu Cao	30.75	—
Yu Xing Cao	31.19	—
Gou Teng	31.71	—
Ge Gen	31.82	—
Wu Zhu Yu	35.87	—
He Huan Pi	36.94	—
Gui Zhi	37.47	—
Xuan Shen	37.96	—
Dan Shen	38.65	—
Huang Lian	39.20	—
Qiang Huo	39.41	—
Qing Hao	40.68	—
Hong Jing Tian	40.93	—
Yu Jin	41.81	—
Huang Bo	42.69	—
Zhi Shi	43.06	—
Hou Po	43.87	—
Fang Feng	44.85	—
Fu Zi	45.14	—
Mang Xiao	47.27	—
Hua Shi	48.40	—
Chen Pi	48.60	—
Hu Zhang	50.75	—

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	8.45
62751	62751	0.00	8.48
62498	62498	0.00	8.71
62752	62752	0.00	8.63

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 2.1-2020-01

Validated substance/substance group **Ju Hua**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002256-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 849 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 *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	25 721
Type B	0	40	0	13 127
Type C	2	0	4	847

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.8560 % (> 99.2683 %)	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
Zi Wan	15.58	—
Gou Qi Zi	16.71	—
Fang Feng	16.86	—
She Gan	16.96	—
Lian Qiao	17.02	—
Du Huo	18.67	—
Hou Po	18.79	—
Chai Hu	19.16	—
Hong Jing Tian	20.37	—
(Shi) Chang Pu	20.65	—
Wu Yao	21.53	—
Lu Gen	21.86	—
Qing Pi	21.96	—
Yi Mu Cao	22.18	—
Cang Zhu	23.04	—
Gui Zhi	23.53	—
Dang Gui	23.98	—
Xiang Fu	24.49	—
Chen Pi	24.78	—
Dang Gui Wei	25.10	—
Bai Mao Gen	25.48	—
Wei Ling Xian	25.86	—
Zhi Mu	26.16	—
Gua Lou	26.36	—
Huang Qi	28.17	—
Mu Dan Pi	28.22	—
Gu Sui Bu	28.37	—
Xie Bai	28.66	—
He Huan Pi	28.94	—
Ze Xie	29.66	—
Qiang Huo	30.16	—
He Shou Wu	30.37	—
Chuan Xiong	31.35	—
Wu Zhu Yu	32.02	—
Chi Shao (Yao)	32.23	—
Xu Duan	32.77	—
Huang Bo	32.88	—
Huang Jing	32.93	—
San Leng	33.10	—
Sha Ren	33.39	—
Mang Xiao	33.89	—
Ma Bian Cao	34.22	—
Ku Shen	34.58	—
Gan Jiang	35.56	—
Jiang Huang	36.49	—
Dan Shen	36.64	—
Ji Li	36.90	—
Xian He Cao	38.17	—
Yin Chen Hao	38.99	—
Sang Ji Shend	39.18	—
Sang Zhi	39.79	—
Ge Gen	40.18	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
(Huai) Niu Xi	40.71	—
Cang Er Zi	40.75	—
Fu Zi	41.09	—
Di Gu Pi	42.06	—
Nü Zhen Zi	42.12	—
Yu Jin	42.32	—
Zhi Zi	42.69	—
Ba Ji Tian	43.15	—
Dang Shen	43.64	—
Jie Geng	43.76	—
Yuan Zhi	44.05	—
Huo Ma Ren	44.16	—
Niu Bang Zi	44.17	—
E Zhu	44.45	—
Xin Yi	44.51	—
Wu Wei Zi	44.59	—
Yin Yang Huo	44.68	—
Sang Bai Pi	45.11	—
Zhi Ke	45.58	—
Shan Yao	46.04	—
Gan Cao	46.47	—
Zhi Shi	46.70	—
Hu Zhang	47.02	—
Long Dan (Cao)	47.02	—
Tu Fu Ling	47.41	—
(Bai) Dou Kou	48.84	—
Pu Gong Ying	49.33	—
Mu Gua	49.71	—
Huang Lian	49.87	—
Bai Shao Yao	50.15	—

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	15.58
62776	62776	0.00	16.28

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 2.1-2020-01

Validated substance/substance group **Jue Ming Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002262-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G059H1097921	62773	40	from supplier

Validation samples

A total of 39 821 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 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.
- 25 761 spectra from a total of 318 batches from further 172 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.

- 79 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	G059H1097921	62773 [†]	20
PhytoComm	Jue Ming Zi	G059H1097921	62774	59

- 13 088 spectra from a total of 319 batches from further 172 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

- 849 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 *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	40	0	25 761
Type B	0	79	0	13 088
Type C	0	0	4	849

The substance/substance group *Jue Ming 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 % (> 99.9705 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.4051 %)
Type C	100.0000 % (> 98.8246 %)	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
Shu Di (Huang)	14.23	—
Ai Ye	20.89	—
(Sheng) Di Huang	26.61	—
Yu Xing Cao	29.22	—
Guang Huo Xiang	33.92	—
Hua Shi	42.44	—
Pu Gong Ying	45.00	—
Jing Jie	52.87	—

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
62773	62773	0.00	14.23

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 2.1-2020-01

Validated substance/substance group **Ku Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60112-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ku Shen	1	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

Validation samples

A total of 39 821 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 *Ku Shen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	1	0	2	850

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.8450 % (> 99.2589 %)	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 *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
Wei Ling Xian	14.10	—
Zhi Zi	14.72	—
(Shi) Chang Pu	18.16	—
Gua Lou	18.55	—
Cang Er Zi	20.77	—
Huo Ma Ren	21.02	—
Ze Xie	22.49	—
Niu Bang Zi	22.58	—
Yi Mu Cao	23.35	—
Xiang Fu	24.15	—
(Huai) Niu Xi	24.35	—
Zi Wan	25.81	—
She Gan	27.14	—
Di Gu Pi	28.13	—
Qing Pi	28.29	—
Gan Cao	29.17	—
Sang Zhi	30.01	—
Lian Qiao	30.04	—
Shan Yao	30.50	—
Tian Hua Fen	32.07	—
Gan Jiang	32.16	—
Shen Qu	32.40	—
Huang Bo	32.48	—
Nü Zhen Zi	32.58	—
Dong Gua Zi	32.59	—
Huang Qi	33.80	—
Dang Gui Wei	34.09	—
Gui Zhi	34.66	—
Jiang Huang	34.88	—
Yuan Zhi	35.42	—
Wu Wei Zi	36.08	—
Mang Xiao	36.29	—
Jin Yin Hua	37.00	—
Wu Yao	37.24	—
Zhi Ke	37.25	—
Chi Shao (Yao)	38.11	—
Mu Gua	39.10	—
Xu Duan	39.22	—
Xie Bai	39.66	—
(Bai) Dou Kou	40.01	—
Gou Qi Zi	40.30	—
Wu Mei	40.70	—
Mu Dan Pi	40.81	—
Qiang Huo	41.10	—
Fu Zi	41.31	—
Bai Mao Gen	42.00	—
Chuan Xiong	42.63	—
Zhi Mu	43.97	—
Jie Geng	44.18	—
Ma Bian Cao	44.21	—
Pu Gong Ying	45.01	—
Dang Shen	45.56	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yu Rou	46.72	—
San Leng	46.82	—
Sha Ren	46.98	—
Zi Su Zi	47.03	—
Bu Gu Zhi	47.41	—
Du Huo	48.02	—
Hou Po	48.41	—
Fang Feng	48.99	—
Chai Hu	49.89	—

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	14.10
62500	62500	0.00	14.72

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 2.1-2020-01

Validated substance/substance group **Lai Fu Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60067-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Lai Fu Zi	1	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

Validation samples

A total of 39 821 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 *Lai Fu Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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

- 850 spectra from 12 *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 *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	0	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

The substance/substance group *Lai 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *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
San Qi	13.40	—
Sheng Jiang	21.15	—
Jie Geng	21.81	—
Di Gu Pi	23.16	—
Mai Men Dong	24.17	—
Sang Zhi	25.45	—
Bai Zhu	25.52	—
Bai Bian Dou	25.85	—
Zhi Mu	26.60	—
Mo Yao	26.74	—
Mang Xiao	26.79	—
Tian Hua Fen	27.09	—
Bai He	27.16	—
Chuan Lian Zi	27.76	—
Gou Qi Zi	29.07	—
She Gan	29.31	—
Qing Pi	29.76	—
Dang Shen	31.43	—
Chai Hu	31.46	—
Long Dan (Cao)	31.72	—
Mu Gua	32.65	—
Dang Gui Wei	35.06	—
Cang Er Zi	35.68	—
Cang Zhu	36.29	—
Du Huo	37.59	—
Mu Dan Pi	37.75	—
Gan Jiang	37.94	—
Ban Xia (Jiang)	37.95	—
Shan Yao	38.34	—
Chuan Xiong	38.81	—
Zi Su Zi	39.13	—
E Zhu	39.27	—
Gui Zhi	39.57	—
Bai Xian Pi	40.20	—
Jiang Huang	40.42	—
San Leng	41.12	—
Dang Gui	41.42	—
Huang Qi	41.50	—
Mai Ya	41.54	—
Ji Li	41.59	—
Lian Zi	42.01	—
Xie Bai	42.84	—
Qiang Huo	43.81	—
Qin Jiao	43.91	—
Tu Fu Ling	44.19	—
Rou Gui	44.29	—
Fu Ling	44.51	—
Da Zao	45.41	—
Gua Lou	46.33	—
Yan Hu Suo	46.80	—
Ba Ji Tian	46.89	—
(Shi) Chang Pu	47.54	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Yi Ren	47.88	—
Ku Shen	47.93	—
Xiang Fu	48.16	—
Zi Wan	48.26	—
Huo Ma Ren	49.15	—
Fang Feng	49.28	—
Bai Shao Yao	49.40	—
Tian Men Dong	49.57	—
Zhu Ru	49.60	—
Yuan Zhi	50.07	—

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	14.36
62492	62492	0.00	13.40

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 2.1-2020-01

Validated substance/substance group **Lian Qiao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60146-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G105H1132721	62231	40	from supplier
PhytoComm	Lian Qiao	G105H1132721	62232	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G105H1132721	62231 [†]	20
PhytoComm	Lian Qiao	G105H1132721	62232 [†]	20
PhytoComm	Lian Qiao	G105H1132821	62337 [†]	20
PhytoComm	Lian Qiao	G105H1132821	62338 [†]	20
PhytoComm	Lian Qiao	G105HS252RW1	62685 [†]	20
PhytoComm	Lian Qiao	G105HS252RW1	62686 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	G105H1132521	1
PhytoComm	Lian Qiao	G105H1132321	1

- 850 spectra from 12 *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 *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	239	1	25 561
Type B	1	119	1	13 046
Type C	0	0	3	850

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 % (> 99.9701 %)	99.5833 % (> 98.3333 %)
Type B	99.9952 % (> 99.9662 %)	99.1667 % (> 96.6667 %)
Type C	100.0000 % (> 98.8257 %)	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
Bai Bian Dou	5.04	—
Yuan Zhi	6.13	—
Shan Yao	6.44	—
Ji Li	6.46	—
Bai Zi Ren	8.92	—
Lian Zi	10.34	—
Zi Wan	10.73	—
Suan Zao Ren	11.17	—
Pi Pa Ye	11.26	—
Ban Xia (Jiang)	11.30	—
Jiao Gu Lan	11.64	—
Chuan Xiong	12.31	—
Shen Qu	12.45	—
Mao Dong Qing	13.18	—
Cang Zhu	13.24	—
Ling Zhi	13.40	—
Fu Zi	13.52	—
Hou Po	14.35	—
San Qi	14.57	—
Jin Yin Hua	14.95	—
Huo Ma Ren	15.28	—
Dan Shen	15.85	—
Ren Dong Teng	16.05	—
Gua Lou	16.84	—
Bai Shao Yao	17.02	—
Chen Pi	17.28	—
Chai Hu	17.57	—
Fu Xiao Mai	17.73	—
Niu Bang Zi	17.87	—
Wu Wei Zi	18.28	—
Zhe Bei Mu	18.30	—
Tai Zi Shen	18.72	—
Tu Fu Ling	18.72	—
Fu Ling	18.83	—
Ze Xie	18.86	—
Sheng Jiang	19.05	—
Rou Gui	19.29	—
Fang Ji	19.38	—
Mu Dan Pi	19.74	—
Shan Yu Rou	20.06	—
Tian Hua Fen	20.67	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
He Huan Pi	20.75	—
Dang Shen	20.85	—
Gou Qi Zi	21.38	—
Gui Zhi	21.69	—
Di Gu Pi	21.71	—
Bai Xian Pi	21.77	—
Yu Zhu	22.13	—
Yan Hu Suo	22.52	—
Dan Dou Chi	22.73	—
Yu Jin	23.15	—
Chi Shao (Yao)	23.51	—
Dang Gui	23.80	—
Zhu Ru	24.24	—
Qing Pi	24.47	—
Chuang Mu Xiang	24.90	—
Yin Yang Huo	25.20	—
Gou Teng	25.29	—
Ye Jiao Teng	25.56	—
Hong Jing Tian	25.89	—
Ji Xue Teng	26.02	—
Tao Ren	27.07	—
Che Qian Zi	27.08	—
Dang Gui Wei	27.35	—
She Gan	27.74	—
Zhi Ke	27.96	—
(Huai) Niu Xi	28.09	—
Yi Yi Ren	28.30	—
Xiang Fu	28.36	—
Du Zhong	28.37	—
Ci Wu Jia	28.62	—
Qiang Huo	28.72	—
Gan Cao	28.80	—
(Shi) Chang Pu	29.84	—
Zhi Mu	30.55	—
Zhi Zi	30.68	—
Ma Huang	30.83	—
Wei Ling Xian	31.54	—
Zi Hua Di Ding	31.65	—
Wu Yao	32.06	—
Ku Shen	32.12	—
Da Zao	32.82	—
Bai Zhu	33.49	—
Bai Mao Gen	34.00	—
Huang Qin	34.09	—
San Leng	34.14	—
Ma Huang Gen	34.68	—
Cang Er Zi	34.69	—
Wang Bu Liu Xing	34.99	—
Sang Zhi	35.22	—
Yi Mu Cao	36.05	—
Huang Lian	36.96	—
Huang Qi	37.47	—
Tian Men Dong	39.39	—
Mang Xiao	39.88	—
Long Yan Rou	40.26	—
Du Huo	41.11	—
Jie Geng	42.28	—
Fo Shou	42.61	—
Ju Hua	44.83	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Bo	44.88	—
Chuan Lian Zi	45.17	—
Sang Ye	45.37	—
Qin Jiao	45.54	—
Jiang Huang	47.10	—
Huang Bai	47.10	—
Fu Shen	47.48	—
Mu Gua	47.93	—
Zi Su Zi	48.50	—
Nü Zhen Zi	49.73	—

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
62231	62231	0.00	10.73
62232	62232	0.00	11.62
62685	62685	0.00	5.72
62338	62338	0.00	17.87
62337	62337	0.00	17.99
62686	62686	0.00	5.04

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 2.1-2020-01

Validated substance/substance group **Lian Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60248-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Lian Zi	2	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	G171HS312QM1	62197	40	from supplier
PhytoComm	Lian Zi	G171HS312QM1	62198	40	from supplier
PhytoComm	Lian Zi	G171HS312SK1	62567	40	from supplier
PhytoComm	Lian Zi	G171HS312SK1	62568	40	from supplier

Validation samples

A total of 39 821 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 *Lian Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G171HS312QM1	62197 [†]	20
PhytoComm	Lian Zi	G171HS312QM1	62198 [†]	20
PhytoComm	Lian Zi	G171HS312SK1	62567 [†]	20
PhytoComm	Lian Zi	G171HS312SK1	62568 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	159	1	25 641
Type B	6	78	2	13 081
Type C	0	0	0	853

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 %)	99.3750 % (> 97.5000 %)
Type B	99.9612 % (> 99.9323 %)	97.5000 % (> 93.7500 %)
Type C	100.0000 % (> 98.8323 %)	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 *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
Bai Bian Dou	4.39	—
Rou Gui	6.51	—
Ban Xia (Jiang)	6.88	—
Gui Zhi	6.89	—
Sheng Jiang	7.07	—
Di Gu Pi	7.13	—
Ji Li	7.67	—
Ling Zhi	8.07	—
He Huan Pi	8.55	—
Tao Ren	9.67	—
Shen Qu	10.63	—
Huo Ma Ren	10.98	—
Ci Wu Jia	11.14	—
Tu Fu Ling	11.18	—
Ren Dong Teng	11.50	—
Bai Shao Yao	11.87	—
Bai Xian Pi	11.93	—
Fo Shou	12.18	—
Fu Xiao Mai	12.32	—
Yuan Zhi	13.02	—
Yi Yi Ren	13.10	—
Bai Zi Ren	13.13	—
Ji Xue Teng	13.42	—
Ye Jiao Teng	13.44	—
Mao Dong Qing	13.75	—
Lian Qiao	13.78	—
Gou Teng	14.04	—
Tai Zi Shen	14.12	—
Zhu Ru	14.41	—
Zhe Bei Mu	14.94	—
Chai Hu	15.00	—
Shan Yao	15.33	—
Fu Ling	15.44	—
Pi Pa Ye	15.73	—
Gua Lou	15.81	—
Chen Pi	15.96	—
Chuan Xiong	16.11	—
Tian Hua Fen	16.29	—
Dan Shen	17.39	—
Fu Zi	17.47	—
Dan Dou Chi	17.64	—
Ma Huang Gen	18.20	—
Yu Jin	18.40	—
Fu Shen	19.38	—
Jiao Gu Lan	20.30	—
Long Yan Rou	20.47	—
Hou Po	20.74	—
San Qi	20.86	—
Tian Men Dong	21.16	—
Cang Zhu	22.13	—
Yu Zhu	22.48	—
Yin Yang Huo	22.88	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Da Zao	23.13	—
Hong Jing Tian	23.41	—
Dang Gui Wei	24.42	—
Che Qian Zi	24.55	—
Suan Zao Ren	25.48	—
Jin Yin Hua	25.66	—
Yan Hu Suo	26.40	—
Ma Huang	26.57	—
Dang Shen	27.42	—
Wang Bu Liu Xing	28.16	—
Fang Ji	28.89	—
Gou Qi Zi	28.91	—
Zhi Mu	29.75	—
Wu Wei Zi	29.94	—
Qiang Huo	30.26	—
Zhi Ke	30.33	—
Dang Gui	31.21	—
Qing Pi	31.31	—
Shan Yu Rou	31.64	—
Mu Dan Pi	34.58	—
Huang Qin	34.96	—
Du Zhong	36.70	—
Mang Xiao	37.00	—
Ze Xie	38.11	—
Sang Zhi	38.18	—
Chuang Mu Xiang	38.33	—
Huang Lian	38.50	—
Chuan Lian Zi	39.35	—
Bai Zhu	39.80	—
Jie Geng	40.79	—
Chi Shao (Yao)	44.13	—
Zi Hua Di Ding	44.56	—
Zi Su Zi	47.66	—
Cang Er Zi	48.69	—
Gan Cao	49.56	—

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
62198	62198	0.00	6.51
62567	62567	0.00	5.15
62197	62197	0.00	7.10
62568	62568	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 2.1-2020-01

Validated substance/substance group **Ling Zhi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60187-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ling Zhi	1	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

Validation samples

A total of 39 821 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 *Ling Zhi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	g109fb2405321	1
PhytoComm	Ling Zhi	G109F-B2405221	1

- 851 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 *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	80	0	25 720
Type B	1	40	0	13 126
Type C	0	0	2	851

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	99.9927 % (> 99.9779 %)	100.0000 % (> 92.5000 %)
Type B	99.9927 % (> 99.9640 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Bai Bian Dou	5.24	—
Ren Dong Teng	5.53	—
Bai Xian Pi	7.19	—
Dan Dou Chi	7.33	—
Yu Jin	7.94	—
Hou Po	8.19	—
Chai Hu	8.60	—
Shen Qu	8.82	—
Lian Zi	9.01	—
Mao Dong Qing	9.18	—
Yuan Zhi	10.08	—
Ji Li	11.23	—
Lian Qiao	11.31	—
Ban Xia (Jiang)	11.61	—
Tu Fu Ling	12.07	—
Yin Yang Huo	13.22	—
Ye Jiao Teng	13.39	—
Shan Yao	14.44	—
Che Qian Zi	14.71	—
He Huan Pi	15.33	—
Yan Hu Suo	15.37	—
Sheng Jiang	15.72	—
Jiao Gu Lan	16.26	—
Dan Shen	16.31	—
Pi Pa Ye	16.48	—
Chuan Xiong	16.67	—
Fu Xiao Mai	16.70	—
Rou Gui	17.01	—
Gou Teng	17.26	—
Huo Ma Ren	17.70	—
Yi Yi Ren	17.98	—
Gui Zhi	18.38	—
Qiang Huo	18.41	—
Fu Zi	18.55	—
Fu Ling	18.81	—
Fang Ji	19.36	—
Tian Hua Fen	19.51	—
Ma Huang Gen	19.72	—
Ji Xue Teng	19.97	—
Qing Pi	20.90	—
Bai Shao Yao	20.91	—
Tao Ren	21.24	—
Du Zhong	21.46	—
Jin Yin Hua	21.70	—
Gua Lou	22.51	—
Bai Zi Ren	22.86	—
Tai Zi Shen	23.07	—
Hong Jing Tian	23.24	—
Zhe Bei Mu	23.31	—
Wu Wei Zi	23.76	—
Chen Pi	24.36	—
Ma Huang	24.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Suan Zao Ren	24.92	—
Di Gu Pi	25.01	—
Mu Dan Pi	25.46	—
Cang Zhu	26.16	—
Ci Wu Jia	27.60	—
Zhu Ru	29.35	—
Zi Hua Di Ding	29.90	—
Zhi Ke	29.98	—
Wang Bu Liu Xing	30.96	—
Huang Lian	31.04	—
San Qi	31.43	—
Gou Qi Zi	32.61	—
Yu Zhu	32.62	—
Da Zao	35.31	—
Long Yan Rou	35.53	—
Fu Shen	35.66	—
Dang Gui	36.57	—
Tian Men Dong	36.63	—
Mang Xiao	37.43	—
Huang Bai	37.44	—
Dang Gui Wei	37.82	—
Dang Shen	38.18	—
Fo Shou	38.18	—
Shan Yu Rou	38.22	—
Chuang Mu Xiang	38.83	—
Huang Qin	41.58	—
Sang Zhi	41.60	—
Ze Xie	43.12	—
Zhi Mu	44.00	—
Jie Geng	45.94	—
Chi Shao (Yao)	46.33	—
Cang Er Zi	46.54	—
Bai Zhu	46.92	—
Pu Gong Ying	47.15	—
Zi Su Zi	48.48	—
Ban Zhi Lian	49.38	—
Xi Xian Cao	49.66	—

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
62454	62454	0.00	5.53
62453	62453	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 2.1-2020-01

Validated substance/substance group **Long Dan (Cao)**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60227-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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)	G112H1705621	62123	40	from supplier
PhytoComm	Long Dan (Cao)	G112H1705621	62124	41	from supplier
PhytoComm	Long Dan (Cao)	G112H1705821	62531	40	from supplier
PhytoComm	Long Dan (Cao)	G112H1705821	62532	40	from supplier

Validation samples

A total of 39 821 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 161 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.
- 25 640 spectra from a total of 317 batches from further 172 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)	G112H1705621	62123 [†]	20
PhytoComm	Long Dan (Cao)	G112H1705621	62124 [†]	20
PhytoComm	Long Dan (Cao)	G112H1705821	62531 [†]	20
PhytoComm	Long Dan (Cao)	G112H1705821	62532 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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)	G112H1705321	3
Phytocomm	Long Dan (Cao)	g112h1705022	1

- 849 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 *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	161	0	25 640
Type B	0	80	0	13 087
Type C	0	0	4	849

The substance/substance group *Long Dan (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 % (> 96.2733 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8246 %)	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 *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
Chuan Xiong	9.54	—
Dang Gui	10.53	—
Dang Gui Wei	13.80	—
Du Huo	14.82	—
Zhi Mu	15.00	—
Bai He	15.08	—
Sang Zhi	15.65	—
Ba Ji Tian	16.10	—
Gou Qi Zi	16.35	—
Di Gu Pi	16.44	—
Jie Geng	17.22	—
Mu Dan Pi	17.64	—
Gui Zhi	18.52	—
She Gan	18.65	—
Mai Ya	18.79	—
Da Zao	19.44	—
San Leng	19.95	—
Mai Men Dong	20.08	—
Chai Hu	20.13	—
Chuan Lian Zi	20.75	—
Mang Xiao	21.22	—
Bai Mao Gen	21.28	—
(Huai) Niu Xi	21.28	—
Fang Feng	21.56	—
Huang Qi	22.75	—
Gan Jiang	23.43	—
Lai Fu Zi	23.46	—
Lu Gen	23.51	—
Bai Shao Yao	24.22	—
Qin Jiao	24.58	—
E Zhu	24.85	—
San Qi	25.40	—
Mu Gua	25.50	—
Sha Shen (Bei)	26.24	—
Chen Pi	26.80	—
Dang Shen	26.97	—
Hou Po	26.99	—
Ji Li	27.67	—
Gua Lou	27.86	—
Bai Bian Dou	27.90	—
Cang Zhu	28.15	—
Zi Wan	28.47	—
Bai Zhu	28.92	—
Mo Yao	30.27	—
Xian He Cao	32.06	—
Tu Fu Ling	33.54	—
He Huan Pi	34.60	—
Lian Qiao	34.78	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sheng Jiang	34.92	—
Xie Bai	34.94	—
Jiang Huang	36.30	—
Hong Jing Tian	36.40	—
Tian Hua Fen	36.41	—
Qing Pi	36.61	—
(Shi) Chang Pu	36.75	—
Zhu Ling	38.38	—
Yan Hu Suo	38.81	—
Ju Hua	40.06	—
Yuan Zhi	40.08	—
Sha Ren	40.12	—
Zi Su Zi	40.44	—
Wu Yao	40.81	—
Yu Jin	41.01	—
Huo Ma Ren	41.32	—
Cang Er Zi	43.82	—
Yi Mu Cao	45.87	—
Zhi Ke	46.33	—
Shan Yao	46.39	—
Ku Shen	47.67	—
Ban Zhi Lian	48.15	—
Chi Shao (Yao)	48.69	—
Xiang Fu	49.25	—
Ban Xia (Jiang)	49.81	—

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	14.15
62123	62123	0.00	9.54
62124	62124	0.00	10.16
62532	62532	0.00	14.21

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 2.1-2020-01

Validated substance/substance group **Long Yan Rou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60463-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Long Yan Rou	2	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	G144HS324RN1	62267	40	from supplier
PhytoComm	Long Yan Rou	G144HS324RN1	62268	40	from supplier
PhytoComm	Long Yan Rou	G144HS324SQ1	62727	40	from supplier
PhytoComm	Long Yan Rou	G144HS324SQ1	62728	40	from supplier

Validation samples

A total of 39 821 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 Yan Rou*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G144HS324RN1	62267 [†]	20
PhytoComm	Long Yan Rou	G144HS324RN1	62268 [†]	20
PhytoComm	Long Yan Rou	G144HS324SQ1	62727 [†]	20
PhytoComm	Long Yan Rou	G144HS324SQ1	62728 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	1	852

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8343 %)	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 *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	7.04	—
Da Zao	10.12	—
Yu Zhu	11.64	—
Gou Qi Zi	11.80	—
Zhi Mu	11.92	—
Gua Lou	12.34	—
Dang Shen	14.77	—
Fo Shou	15.33	—
Ye Jiao Teng	16.42	—
Dang Gui Wei	16.95	—
Di Gu Pi	17.68	—
Tian Men Dong	17.74	—
Chuan Xiong	18.15	—
Bai Shao Yao	18.16	—
Ma Huang	18.28	—
Zhe Bei Mu	18.57	—
Tao Ren	18.98	—
Jie Geng	19.16	—
Cang Zhu	19.56	—
Bai Zhu	20.58	—
Yuan Zhi	20.88	—
Gou Teng	21.37	—
San Qi	21.72	—
Tai Zi Shen	22.04	—
Hong Jing Tian	22.46	—
Chuang Mu Xiang	22.49	—
Tu Fu Ling	22.61	—
Rou Gui	23.01	—
Dang Gui	23.54	—
Zhi Ke	23.67	—
He Huan Pi	23.74	—
Pi Pa Ye	24.05	—
Shen Qu	24.14	—
Gui Zhi	24.28	—
Dan Shen	24.51	—
Ban Xia (Jiang)	24.62	—
Ji Xue Teng	25.15	—
Yin Yang Huo	25.38	—
Ren Dong Teng	25.61	—
Ci Wu Jia	25.64	—
Lian Qiao	25.68	—
Fu Shen	25.69	—
Lian Zi	26.29	—
Huang Qin	26.40	—
Bai Xian Pi	26.57	—
Shan Yu Rou	26.77	—
Jin Yin Hua	26.96	—
Ze Xie	27.10	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gan Cao	27.43	—
Fu Xiao Mai	28.56	—
Shan Yao	28.57	—
Sheng Jiang	28.61	—
Sang Zhi	29.07	—
Ji Li	29.11	—
Bai Bian Dou	29.18	—
San Leng	29.46	—
Ma Huang Gen	29.90	—
Bai Zi Ren	30.13	—
Huo Ma Ren	30.25	—
Jiao Gu Lan	30.34	—
Hou Po	31.22	—
Fu Ling	31.40	—
Fu Zi	31.62	—
Wu Wei Zi	32.49	—
Yi Yi Ren	32.98	—
Huang Qi	33.27	—
Chi Shao (Yao)	33.90	—
Ling Zhi	34.53	—
Dan Dou Chi	34.69	—
Yan Hu Suo	34.84	—
Mai Men Dong	34.85	—
Mao Dong Qing	34.90	—
Zi Hua Di Ding	35.79	—
Chai Hu	36.40	—
Tian Hua Fen	36.65	—
Mu Dan Pi	36.82	—
Zhu Ru	37.93	—
Mang Xiao	39.01	—
Fang Ji	40.18	—
Che Qian Zi	40.55	—
Yu Jin	40.63	—
Huang Lian	42.86	—
Chuan Lian Zi	43.85	—
Suan Zao Ren	44.96	—
Lu Gen	45.20	—
Qiang Huo	45.36	—
Bai He	46.61	—
Mu Gua	47.97	—
Qing Pi	48.00	—
(Shi) Chang Pu	49.16	—
Cang Er Zi	49.36	—

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
62268	62268	0.00	12.95
62267	62267	0.00	13.89
62728	62728	0.00	7.04
62727	62727	0.00	7.22

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 2.1-2020-01

Validated substance/substance group **Lu Gen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60259-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G189H2014821	62485	40	from supplier
PhytoComm	Lu Gen	G189H2014821	62486	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G189H2014821	62485 [†]	20
PhytoComm	Lu Gen	G189H2014821	62486 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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

- 849 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 *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	0	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	4	849

The substance/substance group *Lu 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8246 %)	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
Du Huo	7.75	—
Dang Gui	10.16	—
Gui Zhi	12.63	—
Dang Gui Wei	13.24	—
Chai Hu	14.25	—
Mu Dan Pi	15.02	—
(Shi) Chang Pu	15.51	—
She Gan	15.53	—
Wu Yao	15.97	—
Gou Qi Zi	16.37	—
Bai Mao Gen	16.72	—
Hou Po	16.86	—
Chuan Xiong	16.88	—
Fang Feng	17.33	—
San Leng	18.27	—
Ba Ji Tian	18.75	—
Zi Wan	19.06	—
Ji Li	19.40	—
Sang Zhi	20.47	—
Yi Mu Cao	20.80	—
Zhi Mu	20.91	—
Ju Hua	21.19	—
Xiang Fu	21.47	—
Gan Jiang	21.52	—
Chen Pi	22.02	—
Cang Zhu	22.37	—
Long Dan (Cao)	22.89	—
Gua Lou	23.12	—
Lian Qiao	23.44	—
Di Gu Pi	24.33	—
Hong Jing Tian	24.46	—
Qing Pi	24.75	—
Huang Qi	25.27	—
Xian He Cao	25.31	—
Tu Fu Ling	25.78	—
He Huan Pi	26.37	—
Jiang Huang	26.68	—
E Zhu	26.69	—
Sha Ren	28.59	—
Tian Hua Fen	28.63	—
Mang Xiao	28.64	—
(Huai) Niu Xi	28.85	—
Xie Bai	29.76	—
Chuan Lian Zi	30.41	—
Bai Shao Yao	30.43	—
Jie Geng	31.99	—
Yu Jin	33.23	—
Zhi Ke	33.70	—
Dang Shen	33.74	—
Zi Su Zi	34.19	—
Huo Ma Ren	34.75	—
He Shou Wu	35.17	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yao	35.29	—
(Bai) Dou Kou	35.39	—
Bai Zhu	35.45	—
Yan Hu Suo	35.83	—
Qiang Huo	36.40	—
Huang Jing	36.70	—
Ma Bian Cao	36.96	—
Gu Sui Bu	37.40	—
Lai Fu Zi	38.56	—
Sang Ji Shend	38.90	—
Sha Shen (Bei)	39.01	—
Ku Shen	39.76	—
Ge Gen	39.97	—
Sheng Jiang	40.34	—
Chi Shao (Yao)	41.05	—
Fu Zi	41.12	—
Xu Duan	41.12	—
Ze Xie	41.40	—
Mu Gua	41.40	—
Dan Shen	42.05	—
Huang Lian	42.14	—
Huang Bo	42.31	—
Cang Er Zi	42.93	—
Sang Bai Pi	43.01	—
Bai He	43.11	—
Niu Bang Zi	43.39	—
Xin Yi	43.54	—
Yuan Zhi	43.57	—
Wei Ling Xian	43.65	—
Wu Wei Zi	43.69	—
Qin Jiao	44.06	—
Yin Yang Huo	45.25	—
Da Zao	45.69	—
Wu Zhu Yu	46.25	—
Pu Gong Ying	46.97	—
Ji Xue Teng	48.04	—
Jin Yin Hua	48.50	—
Lian Zi	48.82	—
Yin Chen Hao	48.93	—
Nü Zhen Zi	48.93	—
Zhi Shi	49.55	—
San Qi	49.56	—
Du Zhong	49.89	—

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
62485	62485	0.00	8.09
62486	62486	0.00	7.75

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 2.1-2020-01

Validated substance/substance group **Ma Bian Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60479-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ma Bian Cao; Verbenae officinalis herba

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ma Bian Cao	1	0	1

Second-stage model

For differentiation of the substance/substance group *Ma Bian 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 *Ma Bian Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ma Bian Cao	G248H1085823	62591	40	from supplier
PhytoComm	Ma Bian Cao	G248H1085823	62592	40	from supplier

Validation samples

A total of 39 821 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 Bian Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Bian 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	Ma Bian Cao	G248H1085823	62591 [†]	20
PhytoComm	Ma Bian Cao	G248H1085823	62592 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Ma Bian 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	Ma Bian Cao	G248H1085221	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 *Ma Bian 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 *Ma Bian 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	2	80	0	25 719
Type B	2	37	3	13 125
Type C	0	0	1	852

The substance/substance group *Ma Bian 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	99.9927 % (> 99.9779 %)	100.0000 % (> 92.5000 %)
Type B	99.9855 % (> 99.9567 %)	92.5000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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 *Ma Bian 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
Dan Zhu Ye	4.70	—
Bu Gu Zhi	6.04	—
Gu Sui Bu	14.08	—
Yin Chen Hao	14.20	—
He Shou Wu	14.88	—
Xu Duan	15.92	—
Sha Ren	16.38	—
He Huan Pi	16.59	—
Sang Bai Pi	16.82	—
Xian He Cao	18.92	—
Gui Zhi	19.13	—
Wu Zhu Yu	20.38	—
Jing Jie	21.45	—
Yin Yang Huo	21.69	—
Yi Mu Cao	22.61	—
Sang Ji Shend	23.69	—
Du Zhong	24.62	—
Dan Shen	24.65	—
Yu Jin	25.40	—
Nü Zhen Zi	25.64	—
Fu Zi	28.11	—
Qiang Huo	28.48	—
Chen Pi	28.94	—
Huang Bo	29.92	—
Gou Teng	30.32	—
Jin Qian Cao	31.15	—
(Sheng) Di Huang	34.22	—
Zhi Shi	34.28	—
Jiang Huang	35.23	—
Hou Po	35.67	—
Fang Feng	35.74	—
Bai Hua She She Cao	36.75	—
Sang Ye	36.92	—
Hu Zhang	37.10	—
Xia Ku Cao	37.59	—
Hong Jing Tian	40.11	—
Wu Yao	40.36	—
Ge Gen	40.53	—
Ji Li	40.98	—
Mang Xiao	41.47	—
Pu Gong Ying	41.95	—
Wu Jia Pi	42.25	—
Han Lian Cao	42.71	—
(Bai) Dou Kou	44.08	—
Xin Yi	44.71	—
Huang Lian	45.03	—
Xiang Fu	46.72	—
Xuan Shen	48.53	—
Shan Yao	49.05	—
Cang Er Zi	49.32	—
Ju Hua	49.38	—
Hua Shi	49.79	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Xi Xian Cao	50.03	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ma Bian 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
62591	62591	0.00	4.70
62592	62592	0.00	4.93

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 2.1-2020-01

Validated substance/substance group **Ma Huang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50283-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ma Huang	3	0	1

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	G097H1101623	62147	40	from supplier
PhytoComm	Ma Huang	G097H1101623	62148	40	from supplier
PhytoComm	Ma Huang	G097HS255RT1	62681	40	from supplier
PhytoComm	Ma Huang	G097HS255RT1	62682	40	from supplier
PhytoComm	Ma Huang	G097HS255SL1	62693	40	from supplier
PhytoComm	Ma Huang	G097HS255SL1	62694	40	from supplier

Validation samples

A total of 39 821 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 *Ma Huang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G097H1101623	62147 [†]	20
PhytoComm	Ma Huang	G097H1101623	62148 [†]	20
PhytoComm	Ma Huang	G097HS255RT1	62681 [†]	20
PhytoComm	Ma Huang	G097HS255RT1	62682 [†]	20
PhytoComm	Ma Huang	G097HS255SL1	62693 [†]	20
PhytoComm	Ma Huang	G097HS255SL1	62694 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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 *Ma Huang*.
- 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	Ma Huang	G097H1101522	2

- 851 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	0	2	851

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Bai Shao Yao	12.36	—
Ye Jiao Teng	12.74	—
Zhi Ke	13.03	—
Ji Xue Teng	17.09	—
Hou Po	17.24	—
Jin Yin Hua	17.42	—
Yu Zhu	17.44	—
Pi Pa Ye	17.56	—
Fu Zi	17.96	—
Tai Zi Shen	18.10	—
Chuan Xiong	18.34	—
Gui Zhi	18.55	—
Tao Ren	18.63	—
Lian Qiao	18.81	—
Yin Yang Huo	18.95	—
He Huan Pi	18.99	—
Zi Hua Di Ding	19.58	—
Dan Shen	19.65	—
Huang Lian	20.32	—
Yan Hu Suo	20.33	—
Hong Jing Tian	20.49	—
Gua Lou	20.55	—
Huang Bai	20.68	—
Gou Qi Zi	20.69	—
Huang Qin	20.82	—
Ren Dong Teng	20.93	—
Cang Zhu	21.48	—
Dang Gui	21.83	—
Chen Pi	21.97	—
Tian Hua Fen	21.98	—
Rou Gui	22.07	—
Jiao Gu Lan	22.26	—
Zhe Bei Mu	23.47	—
Bai Xian Pi	23.60	—
Shen Qu	24.45	—
Dang Gui Wei	24.57	—
Bai Bian Dou	25.22	—
Shan Yao	25.45	—
Da Zao	25.55	—
Ci Wu Jia	25.72	—
Di Gu Pi	25.87	—
Gan Cao	25.91	—
Lian Zi	26.01	—
Yuan Zhi	26.10	—
Huo Ma Ren	26.24	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yu Rou	26.62	—
Tu Fu Ling	26.75	—
San Qi	26.79	—
Long Yan Rou	26.88	—
Gou Teng	26.92	—
Wu Wei Zi	27.11	—
Dang Shen	27.12	—
Dan Dou Chi	27.21	—
Sheng Jiang	28.10	—
Mu Dan Pi	28.43	—
Fang Ji	28.75	—
Qiang Huo	29.02	—
Zhi Mu	29.20	—
Ji Li	29.23	—
Ling Zhi	29.73	—
Bai Zi Ren	31.70	—
Ban Xia (Jiang)	32.23	—
Fo Shou	32.56	—
Qing Pi	32.85	—
Fu Shen	33.38	—
Chai Hu	33.40	—
Che Qian Zi	33.49	—
Chuang Mu Xiang	33.92	—
Ma Huang Gen	33.96	—
Yu Jin	33.99	—
Mao Dong Qing	34.46	—
Ban Zhi Lian	34.83	—
Bai Zhu	34.95	—
Chi Shao (Yao)	35.36	—
Pu Gong Ying	36.57	—
Du Zhong	36.67	—
Fu Ling	37.16	—
Sang Ye	37.31	—
Yi Yi Ren	37.85	—
Sang Zhi	39.00	—
Ze Xie	39.53	—
Fu Xiao Mai	39.67	—
Mang Xiao	39.87	—
Wu Zhu Yu	39.91	—
Jie Geng	40.46	—
Tian Men Dong	40.90	—
San Leng	41.15	—
Bai Mao Gen	42.00	—
Wu Yao	42.16	—
Zhi Shi	42.29	—
Zhu Ru	43.01	—
Cang Er Zi	43.40	—
Ge Gen	45.35	—
Xia Ku Cao	45.79	—
Yi Mu Cao	46.44	—
Suan Zao Ren	47.55	—
(Shi) Chang Pu	48.16	—
Huang Bo	49.19	—

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
62681	62681	0.00	13.08
62693	62693	0.00	12.41
62682	62682	0.00	13.03
62148	62148	0.00	39.91
62147	62147	0.00	40.22
62694	62694	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 2.1-2020-01

Validated substance/substance group **Ma Huang Gen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60063-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	79	1	25 720
Type B	3	37	3	13 124
Type C	1	0	0	852

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 %)	98.7500 % (> 95.0000 %)
Type B	99.9855 % (> 99.9567 %)	92.5000 % (> 85.0000 %)
Type C	99.8843 % (> 99.3004 %)	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	—
Gou Teng	6.15	—
Fu Shen	8.53	—
Lian Zi	10.85	—
Ji Xue Teng	11.82	—
Yi Yi Ren	12.57	—
Ban Xia (Jiang)	13.76	—
Sheng Jiang	13.85	—
Fu Ling	14.17	—
Gui Zhi	14.19	—
Zhu Ru	18.57	—
Fo Shou	19.07	—
Fu Xiao Mai	20.06	—
Bai Xian Pi	20.21	—
Ye Jiao Teng	20.48	—
Di Gu Pi	20.60	—
He Huan Pi	21.15	—
Huo Ma Ren	21.64	—
Bai Bian Dou	23.22	—
Ling Zhi	23.45	—
Bai Shao Yao	23.57	—
Tai Zi Shen	24.09	—
Tian Men Dong	24.17	—
Tao Ren	24.92	—
Ci Wu Jia	24.98	—
Ren Dong Teng	25.80	—
Zhe Bei Mu	26.31	—
Da Zao	27.23	—
Chen Pi	27.91	—
Shen Qu	28.02	—
Long Yan Rou	28.07	—
Tu Fu Ling	28.09	—
Bai Zi Ren	28.86	—
Dan Dou Chi	29.21	—
Ji Li	30.96	—
Dan Shen	31.22	—
Yan Hu Suo	32.08	—
Tian Hua Fen	32.72	—
Yu Jin	34.45	—
Lian Qiao	34.63	—
San Qi	34.97	—
Chuan Xiong	34.98	—
Wang Bu Liu Xing	36.21	—
Mang Xiao	36.83	—
Pi Pa Ye	37.00	—
Gua Lou	37.74	—
Yuan Zhi	38.50	—
Zhi Ke	38.52	—
Yin Yang Huo	38.88	—
Zhi Mu	39.47	—
Mao Dong Qing	39.61	—
Jiao Gu Lan	39.71	—
Shan Yao	39.91	—
Ma Huang	41.05	—
Dang Shen	42.04	—
Hong Jing Tian	42.66	—
Yu Zhu	42.70	—
Shan Yu Rou	43.45	—
Fang Ji	44.47	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Qin	44.87	—
Chai Hu	45.32	—
Wu Wei Zi	45.73	—
Fu Zi	46.44	—
Jin Yin Hua	48.08	—
Cang Zhu	49.08	—
Huang Lian	50.03	—

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
62262	62262	0.00	5.97
62261	62261	0.00	3.14

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 2.1-2020-01

Validated substance/substance group **Mai Men Dong**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60024-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Mai Men Dong	3	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	G176H1118621	61945	40	from supplier
PhytoComm	Mai Men Dong	G176H1118621	61946	40	from supplier
PhytoComm	Mai Men Dong	G176H1118622	62113	40	from supplier
PhytoComm	Mai Men Dong	G176H1118622	62114	40	from supplier
PhytoComm	Mai Men Dong	G176HS236SH1	62611	40	from supplier
PhytoComm	Mai Men Dong	G176HS236SH1	62612	40	from supplier

Validation samples

A total of 39 821 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 *Mai Men Dong*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G176H1118621	61945 [†]	20
PhytoComm	Mai Men Dong	G176H1118621	61946 [†]	20
PhytoComm	Mai Men Dong	G176H1118622	62113 [†]	20
PhytoComm	Mai Men Dong	G176H1118622	62114 [†]	20
PhytoComm	Mai Men Dong	G176HS236SH1	62611 [†]	20
PhytoComm	Mai Men Dong	G176HS236SH1	62612 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 849 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	3	1	849

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 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8246 %)	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 *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
Da Zao	11.03	—
Chen Pi	13.82	—
Dang Shen	15.26	—
Tian Men Dong	15.35	—
Zhi Mu	17.18	—
Bai Zhu	18.23	—
Bai He	18.47	—
Jie Geng	20.89	—
Gua Lou	20.98	—
Di Gu Pi	21.25	—
Fo Shou	21.93	—
Ye Jiao Teng	22.96	—
San Leng	24.44	—
Rou Gui	24.90	—
Mang Xiao	26.30	—
Chuan Lian Zi	26.36	—
She Gan	27.09	—
Sang Zhi	27.75	—
Mu Dan Pi	28.53	—
Long Yan Rou	29.74	—
Bai Shao Yao	30.26	—
Yuan Zhi	30.28	—
Fu Xiao Mai	30.33	—
Chuan Xiong	30.94	—
Lai Fu Zi	31.44	—
Chai Hu	31.81	—
Huang Qi	32.10	—
Shan Yu Rou	32.22	—
He Huan Pi	32.27	—
Pi Pa Ye	32.52	—
Tai Zi Shen	34.33	—
Qin Jiao	34.43	—
Dang Gui	34.45	—
San Qi	35.17	—
Yu Zhu	35.19	—
Tao Ren	35.68	—
Gou Teng	35.90	—
Cang Zhu	36.31	—
Tian Hua Fen	36.34	—
Mu Gua	36.35	—
Lian Zi	36.51	—
Fu Ling	36.51	—
Ji Li	36.88	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shen Qu	37.15	—
Fu Shen	37.20	—
Gui Zhi	37.39	—
Chuang Mu Xiang	37.80	—
Bai Xian Pi	38.02	—
Ma Huang	38.33	—
Sheng Jiang	38.54	—
Du Huo	39.27	—
Zhu Ru	39.88	—
Gou Qi Zi	39.94	—
Huang Qin	40.16	—
Ban Xia (Jiang)	40.26	—
Ma Huang Gen	40.28	—
Long Dan (Cao)	40.31	—
Bai Bian Dou	40.82	—
Dang Gui Wei	40.97	—
Ze Xie	41.11	—
Tu Fu Ling	41.34	—
Mai Ya	41.61	—
Zhi Ke	41.82	—
Ji Xue Teng	42.26	—
Shan Yao	43.13	—
Yi Yi Ren	43.81	—
Lian Qiao	44.35	—
Bai Zi Ren	44.80	—
Ba Ji Tian	45.41	—
Ci Wu Jia	46.39	—
Jiao Gu Lan	46.75	—
Ren Dong Teng	46.93	—
Chi Shao (Yao)	47.44	—
Mo Yao	47.57	—
Gan Cao	48.20	—
Zhe Bei Mu	49.64	—

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	12.03
62612	62612	0.00	11.03
62114	62114	0.00	18.70
61946	61946	0.00	18.46
62113	62113	0.00	18.47
61945	61945	0.00	18.23

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 2.1-2020-01

Validated substance/substance group **Mai Ya**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60210-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 850 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	0	1	2	850

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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	9.10	—
Chuan Lian Zi	15.83	—
Chuan Xiong	16.37	—
Di Gu Pi	17.87	—
Long Dan (Cao)	19.47	—
Jie Geng	19.52	—
Sang Zhi	19.84	—
Sheng Jiang	20.21	—
San Qi	20.42	—
Mang Xiao	21.97	—
Zhu Ling	22.57	—
Dang Gui	24.05	—
Mai Men Dong	25.26	—
Ban Zhi Lian	25.52	—
San Leng	26.61	—
Gui Zhi	27.03	—
E Zhu	27.26	—
Da Zao	29.99	—
Mo Yao	30.59	—
Du Huo	30.88	—
Ji Li	30.90	—
Chai Hu	31.08	—
Yan Hu Suo	31.54	—
Mu Dan Pi	31.80	—
Bai Bian Dou	32.22	—
Qin Jiao	32.52	—
Lai Fu Zi	32.66	—
Tian Hua Fen	32.88	—
Bai Zhu	33.29	—
She Gan	33.84	—
Dang Shen	35.32	—
Gan Jiang	35.40	—
Jiang Huang	35.46	—
Dang Gui Wei	35.57	—
Tu Fu Ling	35.88	—
Lian Zi	36.61	—
(Huai) Niu Xi	36.98	—
Mu Gua	37.88	—
Ban Xia (Jiang)	38.84	—
He Huan Pi	39.13	—
Sha Ren	39.26	—
Sha Shen (Bei)	39.46	—
Cang Er Zi	39.57	—
Huang Qi	39.74	—
Ba Ji Tian	39.81	—
Chen Pi	39.98	—
Zhi Mu	40.01	—
Rou Gui	40.20	—
(Bai) Dou Kou	42.16	—
Huo Ma Ren	42.16	—
Cang Zhu	42.43	—
Zi Su Zi	42.95	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yao	43.55	—
Fu Ling	43.90	—
Zhu Ru	44.44	—
Hou Po	44.73	—
Yi Yi Ren	45.15	—
Yu Jin	45.16	—
Xian He Cao	45.48	—
Fang Feng	45.59	—
Lu Gen	46.05	—
Xie Bai	46.51	—
Gou Qi Zi	48.65	—
Qing Pi	49.16	—
(Shi) Chang Pu	49.37	—
Niu Bang Zi	49.82	—
Zi Wan	50.44	—

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	9.20
62536	62536	0.00	9.10

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 2.1-2020-01

Validated substance/substance group **Man Jing Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60025-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Man Jing Zi; Vitis fructus

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	1	852

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8343 %)	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	30.33	—
Da Huang	34.42	—

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	30.37
62502	62502	0.00	30.33

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 2.1-2020-01

Validated substance/substance group **Mang Xiao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50387-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Mang Xiao	1	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

Validation samples

A total of 39 821 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 *Mang Xiao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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 *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.

[†]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
Hua Shi	44.08	—

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	44.08
62470	62470	0.00	44.31

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 2.1-2020-01

Validated substance/substance group **Mao Dong Qing**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60028-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 641
Type B	4	80	0	13 083
Type C	0	0	0	853

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	99.9734 % (> 99.9444 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8323 %)	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	6.98	—
Bai Xian Pi	7.41	—
Yin Yang Huo	7.77	—
Ling Zhi	8.21	—
Yu Jin	9.17	—
Che Qian Zi	10.65	—
Ye Jiao Teng	10.93	—
Hou Po	11.30	—
Fang Ji	11.89	—
Qiang Huo	12.67	—
Xi Xian Cao	13.29	—
Dan Dou Chi	14.95	—
Jiao Gu Lan	15.76	—
Bai Bian Dou	15.76	—
Yuan Zhi	16.40	—
Shen Qu	16.61	—
Ren Dong Teng	16.92	—
Yan Hu Suo	17.46	—
Zi Hua Di Ding	17.67	—
Ji Li	18.18	—
Lian Zi	18.53	—
Mu Dan Pi	18.67	—
Tu Fu Ling	20.12	—
Shan Yao	20.20	—
Huang Bai	20.42	—
Rou Gui	20.47	—
Qing Pi	21.11	—
Pi Pa Ye	22.39	—
Ji Xue Teng	22.44	—
Pu Gong Ying	22.90	—
Chai Hu	23.64	—
Ma Huang Gen	23.74	—
Ma Huang	24.09	—
Wang Bu Liu Xing	24.19	—
Dan Shen	24.29	—
He Huan Pi	24.44	—
Nü Zhen Zi	24.55	—
Wu Mei	24.63	—
Huang Lian	24.73	—
Gou Teng	25.06	—
Chuan Xiong	25.89	—
Zhi Ke	26.27	—
Ban Xia (Jiang)	26.84	—
Lian Qiao	26.89	—
Hong Jing Tian	27.49	—
Jin Yin Hua	28.12	—
Chen Pi	29.47	—
Sang Ye	29.56	—
Wu Wei Zi	29.85	—
Sheng Jiang	30.24	—
Fu Xiao Mai	31.14	—
Bai Shao Yao	31.31	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Yi Ren	31.45	—
Gui Zhi	31.92	—
Ban Zhi Lian	32.31	—
Di Gu Pi	32.77	—
Cang Er Zi	32.98	—
Huo Ma Ren	33.17	—
Xin Yi	33.44	—
Gua Lou	33.57	—
Mang Xiao	34.01	—
Xia Ku Cao	34.02	—
Cang Zhu	35.43	—
Fu Shen	36.79	—
Fu Zi	36.81	—
Tian Hua Fen	37.47	—
E Zhu	37.51	—
Fu Ling	38.22	—
Tao Ren	38.27	—
Jiang Huang	38.32	—
Gou Qi Zi	38.38	—
Tai Zi Shen	39.33	—
Hong Hua	39.47	—
Da Zao	40.41	—
Zhe Bei Mu	41.23	—
Shan Yu Rou	42.09	—
Bai He	43.36	—
Yi Mu Cao	43.85	—
Ze Xie	44.12	—
Sang Zhi	44.68	—
Ci Wu Jia	45.01	—
Chi Shao (Yao)	45.95	—
Zhu Ru	46.30	—
Long Yan Rou	46.79	—
Fo Shou	47.08	—
Jie Geng	47.53	—
Hu Zhang	47.93	—
Suan Zao Ren	48.10	—
Zi Su Zi	48.77	—
Sha Shen (Bei)	49.11	—
Tian Men Dong	49.17	—
Mo Yao	49.34	—
Bai Zi Ren	49.74	—

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
62620	62620	0.00	11.63
62619	62619	0.00	11.71
62417	62417	0.00	6.98
62418	62418	0.00	7.41

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 2.1-2020-01

Validated substance/substance group **Mo Yao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50343-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Mo Yao; Myrrha resina

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Mo Yao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Mo 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 *Mo Yao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Mo Yao	G169H0732823	62479	40	from supplier
PhytoComm	Mo Yao	G169H0732823	62480	40	from supplier

Validation samples

A total of 39 821 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 *Mo Yao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Mo 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	Mo Yao	G169H0732823	62479 [†]	20
PhytoComm	Mo Yao	G169H0732823	62480 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Mo 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
Caelo	Mo Yao	G169H0732222	1
Phytocomm	Mo Yao	G1690732222	1
PhytoComm	Mo Yao	G169H0732222	1

- 850 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 *Mo 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 *Mo 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	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

The substance/substance group *Mo 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *Mo 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
Bai Bian Dou	12.31	—
Mang Xiao	25.79	—
E Zhu	26.08	—
San Qi	26.92	—
Cang Er Zi	27.22	—
Lai Fu Zi	28.56	—
Zi Su Zi	30.25	—
Jiang Huang	30.96	—
Sang Zhi	35.18	—
Chuan Lian Zi	36.39	—
She Gan	36.68	—
Sheng Jiang	36.72	—
Bai He	38.53	—
Wang Bu Liu Xing	38.65	—
Chai Hu	39.16	—
Shan Yao	40.32	—
Ji Li	40.58	—
Mai Men Dong	40.65	—
Gou Qi Zi	40.92	—
Xie Bai	41.34	—
Di Gu Pi	41.37	—
Long Dan (Cao)	41.60	—
Da Zao	42.77	—
Ban Xia (Jiang)	43.37	—
Gui Zhi	44.12	—
Mai Ya	44.49	—
Chuan Xiong	45.04	—
He Huan Pi	45.61	—
Zhi Mu	46.18	—
Dong Gua Zi	46.61	—
Niu Bang Zi	46.62	—
Huo Ma Ren	46.72	—
Jie Geng	46.83	—
Ba Ji Tian	47.43	—
Sha Ren	48.37	—
Cang Zhu	49.74	—
Tian Hua Fen	49.75	—
Du Huo	49.87	—
Fu Ling	49.89	—
Suan Zao Ren	50.10	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Mo 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
62480	62480	0.00	12.31
62479	62479	0.00	12.74

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 2.1-2020-01

Validated substance/substance group **Mu Dan Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60113-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Mu Dan Pi	2	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	G166HS129RT1	62433	40	from supplier
PhytoComm	Mu Dan Pi	G166HS129RT1	62434	40	from supplier
PhytoComm	Mu Dan Pi	G166H0741821	62475	40	from supplier
PhytoComm	Mu Dan Pi	G166H0741821	62476	40	from supplier

Validation samples

A total of 39 821 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 Dan Pi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G166HS129RT1	62433 [†]	20
PhytoComm	Mu Dan Pi	G166HS129RT1	62434 [†]	20
PhytoComm	Mu Dan Pi	G166H0741821	62475 [†]	20
PhytoComm	Mu Dan Pi	G166H0741821	62476 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	g166j074122	1
Phytocomm	Mu Dan Pi	G166H0741421	4
PhytoComm	Mu Dan Pi	G166H0741321	1

- 842 spectra from 12 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	76	4	13 087
Type C	0	0	11	842

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	95.0000 % (> 91.2500 %)
Type C	100.0000 % (> 98.8225 %)	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 *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
Dan Shen	5.62	—
Zi Hua Di Ding	7.71	—
Du Huo	10.33	—
Jin Yin Hua	11.56	—
Chuan Xiong	11.59	—
Hou Po	12.00	—
Long Dan (Cao)	12.47	—
Dang Gui	12.47	—
Yin Yang Huo	14.55	—
Jiao Gu Lan	14.57	—
Qiang Huo	14.80	—
Lu Gen	14.89	—
Huang Bai	15.35	—
Ye Jiao Teng	15.71	—
She Gan	15.96	—
Hong Jing Tian	15.97	—
Zhi Ke	15.99	—
Chai Hu	16.15	—
San Leng	16.24	—
Shan Yao	16.34	—
Bai Xian Pi	17.19	—
Fang Ji	17.30	—
Dan Dou Chi	17.31	—
Yuan Zhi	18.35	—
Mao Dong Qing	18.48	—
Pi Pa Ye	18.55	—
Bai Mao Gen	18.63	—
Bai Shao Yao	18.65	—
Wu Wei Zi	18.65	—
Dang Gui Wei	18.85	—
Huang Lian	18.95	—
Qing Pi	19.32	—
Ren Dong Teng	19.39	—
Ji Xue Teng	19.49	—
Gou Qi Zi	19.59	—
Di Gu Pi	19.82	—
Yu Jin	20.36	—
Ma Huang	20.44	—
Du Zhong	20.53	—
Fu Zi	20.74	—
Sang Zhi	20.88	—
Bai Bian Dou	20.98	—
Gua Lou	21.01	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Lian Zi	21.15	—
Gui Zhi	21.32	—
Ba Ji Tian	23.25	—
Fang Feng	23.55	—
Yan Hu Suo	23.72	—
Che Qian Zi	23.74	—
Huang Qi	23.80	—
Shen Qu	23.97	—
Pu Gong Ying	24.16	—
Zhi Mu	24.45	—
Ji Li	25.27	—
Lian Qiao	25.33	—
Ling Zhi	25.50	—
He Huan Pi	25.91	—
Cang Zhu	25.96	—
Tu Fu Ling	26.17	—
(Shi) Chang Pu	26.57	—
Zi Wan	26.58	—
Shan Yu Rou	26.96	—
Gou Teng	27.82	—
Chen Pi	28.15	—
Gan Jiang	28.16	—
Mang Xiao	28.51	—
Jie Geng	28.87	—
(Huai) Niu Xi	28.93	—
Wu Yao	29.10	—
Da Zao	29.23	—
Ze Xie	29.34	—
Chi Shao (Yao)	29.73	—
Rou Gui	29.97	—
Cang Er Zi	30.04	—
Xian He Cao	30.33	—
E Zhu	31.04	—
Sang Ye	31.31	—
Bai Zhu	31.43	—
Tao Ren	31.75	—
Zhe Bei Mu	32.43	—
Huo Ma Ren	33.16	—
Qin Jiao	33.61	—
Sha Shen (Bei)	33.61	—
Gan Cao	34.24	—
Dang Shen	34.25	—
Tian Hua Fen	34.43	—
Long Yan Rou	34.56	—
Huang Qin	34.95	—
Ban Zhi Lian	35.02	—
Yi Mu Cao	35.07	—
Ju Hua	35.67	—
Mu Gua	35.68	—
Ma Huang Gen	35.78	—
San Qi	35.90	—
Chuang Mu Xiang	36.14	—
Jiang Huang	36.73	—
Sheng Jiang	36.83	—
Hong Hua	37.64	—
Chuan Lian Zi	37.79	—
Xiang Fu	39.10	—
Xi Xian Cao	39.11	—
Ban Xia (Jiang)	39.75	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tai Zi Shen	39.88	—
Fu Xiao Mai	39.97	—
Sha Ren	40.83	—
Ci Wu Jia	40.83	—
Xie Bai	41.20	—
Fu Ling	41.63	—
Suan Zao Ren	42.27	—
Bai He	42.31	—
Xia Ku Cao	43.28	—
Lai Fu Zi	43.49	—
Yu Zhu	43.73	—
Huang Jing	44.09	—
Zi Su Zi	44.57	—
Niu Bang Zi	45.05	—
Bai Zi Ren	45.43	—
Yi Yi Ren	45.49	—
Fu Shen	46.29	—
Gu Sui Bu	46.68	—
Mai Ya	47.33	—
Fo Shou	47.36	—
He Shou Wu	47.74	—
Sang Ji Shend	47.89	—
Zhu Ru	48.50	—
Mai Men Dong	48.99	—
Wu Mei	49.08	—
Zhi Shi	49.38	—
Ku Shen	49.93	—
(Bai) Dou Kou	50.09	—

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
62433	62433	0.00	5.81
62476	62476	0.00	10.33
62475	62475	0.00	10.47
62434	62434	0.00	5.62

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 2.1-2020-01

Validated substance/substance group **Mu Gua**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60236-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Mu Gua	1	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	G063H0424621	62085	40	from supplier
PhytoComm	Mu Gua	G063H0424621	62086	40	from supplier

Validation samples

A total of 39 821 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 Gua*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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	G063H0424621	62085 [†]	20
PhytoComm	Mu Gua	G063H0424621	62086 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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	G063H0424521	4
PhytoComm	Mu Gua	G063H0424221	1

- 847 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	6	0	6	841

The substance/substance group *Mu Gua* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.5770 % (> 98.9887 %)	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 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	9.22	—
Jie Geng	11.62	—
Chi Shao (Yao)	15.82	—
Bai Shao Yao	18.23	—
San Leng	19.20	—
She Gan	19.29	—
(Huai) Niu Xi	19.69	—
Gou Qi Zi	20.02	—
Bai Zhu	22.86	—
Qin Jiao	23.01	—
Dang Gui Wei	23.32	—
Yuan Zhi	24.07	—
Di Gu Pi	25.10	—
Dang Shen	25.68	—
Sang Zhi	25.83	—
Zi Wan	28.04	—
Chuan Xiong	29.43	—
Bai Mao Gen	29.64	—
Ze Xie	30.54	—
Cang Zhu	32.06	—
Lian Qiao	32.22	—
Gui Zhi	33.70	—
Mu Dan Pi	34.65	—
(Shi) Chang Pu	34.77	—
Long Dan (Cao)	34.85	—
Zhi Mu	35.09	—
Mang Xiao	35.16	—
Gua Lou	35.36	—
Qing Pi	35.48	—
Chai Hu	35.69	—
Chen Pi	36.36	—
Du Huo	36.41	—
Gan Cao	36.55	—
Zhi Ke	36.67	—
Lu Gen	37.70	—
Lai Fu Zi	38.27	—
Xiang Fu	39.33	—
Shan Yu Rou	39.78	—
Gan Jiang	42.60	—
Bai He	42.99	—
Sheng Jiang	43.62	—
San Qi	44.67	—
Long Yan Rou	44.71	—
Da Zao	45.14	—
Dang Gui	46.13	—
Huang Qin	46.18	—
Tian Men Dong	46.27	—
Lian Zi	46.54	—
Fang Feng	46.95	—
Bai Xian Pi	47.16	—
Tu Fu Ling	47.63	—
Mai Men Dong	48.69	—

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
62085	62085	0.00	9.26
62086	62086	0.00	9.22

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 2.1-2020-01

Validated substance/substance group **Niu Bang Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002261-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Niu Bang Zi	1	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

Validation samples

A total of 39 821 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 *Niu Bang Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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	G026H0433521	3
PhytoComm	Niu Bang Zi	G026H0433321	2

- 845 spectra from 12 *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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	8	845

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8230 %)	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 *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
Gua Lou	16.93	—
Qing Pi	17.16	—
Zhi Zi	18.89	—
Ku Shen	20.07	—
Wei Ling Xian	20.76	—
Lian Qiao	21.13	—
She Gan	22.48	—
Zi Wan	25.26	—
Sang Zhi	28.53	—
Ze Xie	28.60	—
Di Gu Pi	29.34	—
Gou Qi Zi	30.48	—
Gui Zhi	30.56	—
Cang Er Zi	31.92	—
Huo Ma Ren	32.46	—
(Huai) Niu Xi	32.92	—
Jiang Huang	33.48	—
Dang Gui Wei	33.98	—
Xie Bai	34.77	—
Cang Zhu	34.96	—
Du Huo	35.14	—
Mang Xiao	35.54	—
Xiang Fu	35.84	—
Yi Mu Cao	36.39	—
Shan Yao	36.60	—
Chai Hu	36.69	—
Wu Yao	36.98	—
Chi Shao (Yao)	37.60	—
Hou Po	37.92	—
Zhi Ke	38.42	—
(Shi) Chang Pu	39.03	—
Huang Bo	39.13	—
Dong Gua Zi	39.40	—
Gan Jiang	39.58	—
Tian Hua Fen	40.03	—
Chuan Xiong	40.70	—
Nü Zhen Zi	40.81	—
Huang Qi	41.67	—
Wu Wei Zi	41.72	—
Jie Geng	42.38	—
Bai Mao Gen	42.39	—
Fang Feng	42.40	—
Gan Cao	42.55	—
Hong Jing Tian	42.69	—
Ju Hua	43.63	—
Yuan Zhi	43.76	—
(Bai) Dou Kou	44.77	—
Mu Dan Pi	44.86	—
Sha Ren	45.11	—
Shen Qu	45.32	—
Yan Hu Suo	46.42	—
San Leng	46.62	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Qiang Huo	46.70	—
Jin Yin Hua	46.84	—
Zi Su Zi	47.02	—
Ji Li	47.17	—
Xian He Cao	47.56	—
Chen Pi	47.60	—
Zhi Mu	48.86	—
Dang Gui	49.23	—
Dang Shen	49.72	—
E Zhu	50.43	—

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	16.93
62766	62766	0.00	17.16

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 2.1-2020-01

Validated substance/substance group **Nü Zhen Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60065-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Nü Zhen Zi	3	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	G138H0345721	62213	40	from supplier
PhytoComm	Nü Zhen Zi	G138H0345721	62214	40	from supplier
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 39 821 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 *Nü Zhen Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G138H0345721	62213 [†]	20
PhytoComm	Nü Zhen Zi	G138H0345721	62214 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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

- 850 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	0	3	850

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *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
Yin Chen Hao	10.01	—
Qiang Huo	11.63	—
Xu Duan	12.90	—
Dan Shen	16.21	—
Yin Yang Huo	18.11	—
Xi Xian Cao	18.46	—
Sang Ye	18.58	—
Fang Ji	18.70	—
Ma Bian Cao	18.79	—
Xin Yi	18.97	—
Dan Zhu Ye	19.00	—
Yi Mu Cao	19.03	—
Shan Yao	19.56	—
Gou Teng	19.95	—
Huang Bo	20.91	—
Hou Po	21.49	—
Wu Zhu Yu	22.78	—
Jing Jie	22.93	—
Pu Gong Ying	22.94	—
Du Zhong	23.22	—
Sang Ji Shend	23.95	—
Hong Jing Tian	24.12	—
Cang Er Zi	24.18	—
Yu Jin	24.88	—
Mao Dong Qing	24.89	—
Qing Pi	26.36	—
Ling Zhi	26.47	—
Sha Ren	26.71	—
Xia Ku Cao	26.91	—
Wei Ling Xian	27.22	—
Tu Fu Ling	27.83	—
Huang Lian	28.57	—
Gui Zhi	28.64	—
Bu Gu Zhi	28.66	—
Jiang Huang	29.09	—
Hu Zhang	29.40	—
Wu Wei Zi	29.68	—
Jiao Gu Lan	29.97	—
Fu Zi	30.21	—
Zhi Ke	30.34	—
He Huan Pi	30.53	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Mei	31.51	—
Xian He Cao	31.74	—
Bai Xian Pi	32.80	—
Sang Bai Pi	33.06	—
Di Gu Pi	33.40	—
Gu Sui Bu	33.60	—
Hong Hua	34.23	—
Pi Pa Ye	34.71	—
Wang Bu Liu Xing	34.94	—
Han Lian Cao	35.34	—
Dan Dou Chi	35.74	—
Ye Jiao Teng	35.87	—
Shen Qu	36.12	—
Bai Hua She She Cao	36.39	—
He Shou Wu	37.32	—
Ma Huang	37.72	—
Jin Yin Hua	37.84	—
Fang Feng	38.84	—
Huang Bai	38.86	—
Zi Hua Di Ding	39.58	—
Yan Hu Suo	39.82	—
Wu Yao	40.29	—
Chuan Xiong	40.29	—
Bai Bian Dou	40.85	—
Zhi Shi	41.02	—
Ku Shen	41.39	—
Ji Xue Teng	41.51	—
Chai Hu	41.81	—
Ji Li	41.83	—
Ge Gen	42.03	—
Shan Yu Rou	42.04	—
Bai Shao Yao	42.33	—
Mang Xiao	42.47	—
Wu Jia Pi	42.82	—
Qing Hao	44.01	—
Lian Qiao	44.21	—
Ren Dong Teng	44.40	—
Chen Pi	44.75	—
(Sheng) Di Huang	45.67	—
Ze Xie	45.90	—
Lian Zi	46.50	—
Rou Cong Rong	46.92	—
E Zhu	46.94	—
Jin Qian Cao	48.96	—
Sang Zhi	49.50	—
Cang Zhu	49.94	—

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
62569	62569	0.00	18.56
62213	62213	0.00	10.24
62214	62214	0.00	10.01
62570	62570	0.00	18.46
62343	62343	0.00	11.63
62344	62344	0.00	12.10

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 2.1-2020-01

Validated substance/substance group **Pi Pa Ye**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60201-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	3	78	2	25 718
Type B	4	33	7	13 123
Type C	0	0	1	852

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	99.9927 % (> 99.9779 %)	97.5000 % (> 93.7500 %)
Type B	99.9830 % (> 99.9543 %)	82.5000 % (> 75.0000 %)
Type C	100.0000 % (> 98.8343 %)	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	4.21	—
Lian Qiao	5.91	—
Fu Zi	7.00	—
Ren Dong Teng	7.52	—
Ji Li	8.49	—
Chuan Xiong	8.97	—
Gua Lou	9.13	—
Bai Bian Dou	9.32	—
Bai Zi Ren	9.37	—
Shen Qu	9.62	—
Huo Ma Ren	9.85	—
Tian Hua Fen	10.04	—
Cang Zhu	10.42	—
Jin Yin Hua	11.29	—
Chen Pi	11.31	—
Bai Shao Yao	11.39	—
San Qi	11.73	—
Zhe Bei Mu	12.56	—
Yuan Zhi	13.16	—
Hou Po	13.44	—
Tai Zi Shen	13.52	—
Jiao Gu Lan	13.86	—
Lian Zi	14.13	—
He Huan Pi	15.05	—
Dan Shen	15.10	—
Ban Xia (Jiang)	15.87	—
Fu Ling	16.37	—
Mao Dong Qing	16.50	—
Ling Zhi	16.58	—
Fu Xiao Mai	17.48	—
Gou Qi Zi	17.51	—
Ye Jiao Teng	17.62	—
Hong Jing Tian	17.90	—
Ci Wu Jia	18.01	—
Dang Gui Wei	18.23	—
Tao Ren	18.28	—
Ji Xue Teng	18.44	—
Zhu Ru	18.45	—
Rou Gui	18.51	—
Di Gu Pi	18.55	—
Gou Teng	18.85	—
Yu Zhu	19.07	—
Tu Fu Ling	19.20	—
Yan Hu Suo	19.27	—
Dang Gui	19.51	—
Dang Shen	19.57	—
Che Qian Zi	19.77	—
Gui Zhi	19.81	—
Sheng Jiang	19.93	—
Shan Yu Rou	20.00	—
Wu Wei Zi	20.08	—
Chai Hu	20.31	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yin Yang Huo	20.45	—
Fang Ji	21.38	—
Da Zao	21.45	—
Bai Xian Pi	21.70	—
Dan Dou Chi	22.15	—
Ma Huang	22.42	—
Yi Yi Ren	23.10	—
Yu Jin	23.53	—
Mu Dan Pi	23.66	—
Zhi Mu	23.72	—
Qing Pi	23.84	—
Fo Shou	24.33	—
Suan Zao Ren	24.85	—
Chuang Mu Xiang	25.44	—
Zhi Ke	26.51	—
Qiang Huo	26.80	—
Long Yan Rou	27.20	—
Ma Huang Gen	27.92	—
Huang Qin	29.32	—
Gan Cao	29.75	—
Tian Men Dong	29.81	—
Zi Hua Di Ding	31.89	—
Bai Zhu	32.50	—
Sang Zhi	32.83	—
Huang Lian	33.24	—
Du Zhong	33.40	—
Ze Xie	35.19	—
Wang Bu Liu Xing	35.29	—
Chi Shao (Yao)	36.95	—
Mang Xiao	38.35	—
Cang Er Zi	38.38	—
Jie Geng	39.82	—
Fu Shen	39.98	—
San Leng	43.58	—
Chuan Lian Zi	44.42	—
Sang Ye	46.95	—
Huang Bai	48.36	—
Bai Mao Gen	48.90	—
Hong Hua	48.91	—
(Shi) Chang Pu	49.03	—
Zi Su Zi	49.28	—

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	4.21
62452	62452	0.00	4.22

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 2.1-2020-01

Validated substance/substance group **Pu Gong Ying**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50358-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Pu Gong Ying	3	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	G239H1315623	62119	40	from supplier
PhytoComm	Pu Gong Ying	G239H1315623	62120	40	from supplier
PhytoComm	Pu Gong Ying	G239H1315623	62163	40	from supplier
PhytoComm	Pu Gong Ying	G239H1315623	62164	40	from supplier
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 39 821 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 *Pu Gong Ying*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 481 spectra from a total of 316 batches from further 172 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 *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	G239H1315623	62119 [†]	20
PhytoComm	Pu Gong Ying	G239H1315623	62120 [†]	20
PhytoComm	Pu Gong Ying	G239H1315623	62163 [†]	20
PhytoComm	Pu Gong Ying	G239H1315623	62164 [†]	20

continued on the next page

continued from previous page

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

- 13 007 spectra from a total of 317 batches from further 172 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

- 847 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 *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	320	0	25 481
Type B	0	159	1	13 007
Type C	3	0	6	844

The substance/substance group *Pu Gong Ying* 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 % (> 99.9701 %)	100.0000 % (> 98.1250 %)
Type B	100.0000 % (> 99.9421 %)	99.3750 % (> 97.5000 %)
Type C	99.7121 % (> 99.1238 %)	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 *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
Huang Lian	7.65	—
Sang Ye	7.80	—
Xin Yi	8.71	—
Tu Fu Ling	9.22	—
Xi Xian Cao	11.45	—
Xia Ku Cao	13.99	—
Shan Yao	14.50	—
Nü Zhen Zi	16.35	—
Jin Qian Cao	17.79	—
Gui Zhi	18.17	—
Dan Shen	18.66	—
Yin Chen Hao	18.82	—
Hu Zhang	19.35	—
Jiang Huang	20.06	—
Fu Zi	20.69	—
Han Lian Cao	21.85	—
Zhi Ke	21.99	—
Gou Teng	22.36	—
Du Zhong	22.61	—
Cang Er Zi	23.34	—
Mao Dong Qing	24.64	—
Jing Jie	25.37	—
Qiang Huo	25.43	—
Wu Jia Pi	25.45	—
(Sheng) Di Huang	25.74	—
Yin Yang Huo	26.51	—
Dan Zhu Ye	27.28	—
Xuan Shen	27.37	—
Jiao Gu Lan	27.72	—
Qing Hao	28.31	—
Fang Ji	28.60	—
Xu Duan	29.05	—
Sha Ren	29.10	—
Hong Jing Tian	29.39	—
Wu Mei	29.90	—
Yu Jin	29.93	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yi Mu Cao	30.46	—
Sang Ji Shend	30.63	—
Hou Po	30.64	—
Ye Jiao Teng	31.14	—
Yan Hu Suo	31.85	—
Bai Hua She She Cao	32.15	—
Qing Pi	32.25	—
Wu Wei Zi	32.34	—
Guang Huo Xiang	32.73	—
Xiang Fu	33.75	—
Ma Huang	34.11	—
Yu Xing Cao	34.38	—
Wu Yao	34.70	—
Zhi Shi	35.54	—
Bai Xian Pi	35.93	—
Bu Gu Zhi	36.16	—
Ze Xie	36.19	—
Huang Bai	36.42	—
Shen Qu	36.57	—
Di Gu Pi	36.76	—
Huang Bo	36.89	—
Jin Yin Hua	37.26	—
He Huan Pi	37.85	—
Pi Pa Ye	38.15	—
(Shi) Chang Pu	38.45	—
Mang Xiao	38.60	—
Dan Dou Chi	38.95	—
Ge Gen	39.00	—
Xian He Cao	39.00	—
Che Qian Zi	39.63	—
Mu Dan Pi	39.63	—
Zi Hua Di Ding	39.78	—
Wei Ling Xian	40.08	—
Chai Hu	40.15	—
Wang Bu Liu Xing	40.66	—
Gou Qi Zi	40.76	—
Lu Gen	41.27	—
Bai Mao Gen	41.95	—
E Zhu	42.00	—
Gua Lou	42.02	—
Ji Li	42.16	—
Chi Shao (Yao)	42.45	—
Ji Xue Teng	42.64	—
Bai Shao Yao	42.87	—
Chuan Xiong	43.15	—
Hong Hua	43.63	—
Gu Sui Bu	43.74	—
Ling Zhi	44.04	—
Wu Zhu Yu	44.46	—
Lian Zi	44.70	—
Sang Zhi	45.23	—
Shan Yu Rou	45.92	—
Ma Bian Cao	46.47	—
Lian Qiao	46.90	—
(Bai) Dou Kou	47.03	—
Du Huo	47.15	—
Hua Shi	47.47	—
Ren Dong Teng	47.51	—
Sang Bai Pi	47.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ai Ye	47.64	—
Shu Di (Huang)	47.73	—
He Shou Wu	47.85	—
She Gan	47.96	—
Zi Wan	48.73	—
Rou Cong Rong	49.82	—

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
62119	62119	0.00	18.88
62120	62120	0.00	18.82
62264	62264	0.00	7.65
62163	62163	0.00	17.79
62636	62636	0.00	8.05
62263	62263	0.00	8.05
62164	62164	0.00	17.94
62635	62635	0.00	7.80

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 2.1-2020-01

Validated substance/substance group **Qiang Huo**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60109-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Qiang Huo	2	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	G172H0840721	62361	40	from supplier
PhytoComm	Qiang Huo	G172H0840721	62362	40	from supplier
PhytoComm	Qiang Huo	G172HS154SL1	62719	40	from supplier
PhytoComm	Qiang Huo	G172HS154SL1	62720	40	from supplier

Validation samples

A total of 39 821 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 *Qiang Huo*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G172H0840721	62361 [†]	20
PhytoComm	Qiang Huo	G172H0840721	62362 [†]	20
PhytoComm	Qiang Huo	G172HS154SL1	62719 [†]	20
PhytoComm	Qiang Huo	G172HS154SL1	62720 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	g172h0841321	1
PhytoComm	Qiang Huo	G172H0840321	1

- 850 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 *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	160	0	25 641
Type B	0	78	2	13 087
Type C	1	1	2	849

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	97.5000 % (> 93.7500 %)
Type C	99.9225 % (> 99.3353 %)	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 *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	6.00	—
Ling Zhi	7.62	—
Yu Jin	7.62	—
Hou Po	8.04	—
Du Zhong	9.02	—
Bai Xian Pi	9.64	—
Yin Yang Huo	10.25	—
Dan Dou Chi	10.54	—
Fang Ji	11.20	—
Qing Pi	11.32	—
Dan Shen	11.53	—
Nü Zhen Zi	11.89	—
Bai Bian Dou	11.94	—
Zi Hua Di Ding	12.93	—
Yi Mu Cao	13.38	—
Chai Hu	13.80	—
Huang Bo	15.06	—
Shen Qu	15.12	—
Ye Jiao Teng	15.56	—
Lian Zi	15.77	—
Xu Duan	15.83	—
Jiao Gu Lan	15.83	—
Mu Dan Pi	15.88	—
Ji Li	16.67	—
Shan Yao	17.22	—
Yan Hu Suo	17.40	—
Pi Pa Ye	17.43	—
Yuan Zhi	17.81	—
Yin Chen Hao	18.16	—
Ma Bian Cao	18.56	—
Hong Jing Tian	18.62	—
Ren Dong Teng	18.86	—
He Huan Pi	18.99	—
Che Qian Zi	19.12	—
Gui Zhi	20.19	—
Lian Qiao	20.22	—
Chuan Xiong	20.68	—
Huang Bai	21.68	—
Ban Xia (Jiang)	21.98	—
Jin Yin Hua	22.14	—
Wei Ling Xian	22.19	—
Wu Zhu Yu	22.83	—
Ma Huang	22.90	—
Fu Zi	23.22	—
Huang Lian	23.42	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Wei Zi	23.91	—
Ji Xue Teng	24.25	—
Sang Ji Shend	24.27	—
Bai Shao Yao	24.96	—
Gu Sui Bu	25.64	—
Gua Lou	25.97	—
Sha Ren	26.25	—
Zhi Ke	26.26	—
Di Gu Pi	26.29	—
Gou Teng	26.36	—
Tu Fu Ling	26.49	—
Wu Yao	26.80	—
Pu Gong Ying	27.05	—
Sheng Jiang	27.20	—
Rou Gui	27.42	—
Chen Pi	27.46	—
Bu Gu Zhi	27.72	—
Ma Huang Gen	27.73	—
Dan Zhu Ye	27.86	—
Cang Zhu	28.09	—
Fu Xiao Mai	29.08	—
Xi Xian Cao	29.28	—
Sang Bai Pi	29.54	—
Gou Qi Zi	29.66	—
Jiang Huang	30.48	—
Fang Feng	30.75	—
Hu Zhang	31.65	—
Xian He Cao	31.91	—
Yi Yi Ren	32.19	—
Wang Bu Liu Xing	32.68	—
Ju Hua	32.76	—
Huo Ma Ren	33.56	—
Xiang Fu	33.89	—
Sang Ye	34.40	—
Tao Ren	35.05	—
Jing Jie	35.14	—
Mang Xiao	35.87	—
Zhe Bei Mu	35.96	—
Ku Shen	36.32	—
Ze Xie	36.36	—
Tian Hua Fen	36.45	—
Fu Ling	36.72	—
Cang Er Zi	37.91	—
Xia Ku Cao	38.00	—
Da Zao	38.46	—
Shan Yu Rou	38.60	—
Ban Zhi Lian	38.79	—
Sang Zhi	38.97	—
Suan Zao Ren	39.61	—
Wu Mei	39.72	—
Tai Zi Shen	39.96	—
Bai Zi Ren	41.08	—
Ge Gen	41.19	—
Jie Geng	41.42	—
Chi Shao (Yao)	41.52	—
(Shi) Chang Pu	41.61	—
He Shou Wu	42.11	—
San Qi	42.32	—
Ci Wu Jia	42.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Jing	43.61	—
Zi Su Zi	43.87	—
(Bai) Dou Kou	44.15	—
Chuang Mu Xiang	44.34	—
Hong Hua	44.41	—
E Zhu	44.67	—
Bai Zhu	44.87	—
Zhu Ru	44.88	—
Rou Cong Rong	44.89	—
Long Yan Rou	45.07	—
Xie Bai	45.08	—
Fu Shen	45.22	—
Huang Qin	45.46	—
Zhi Mu	45.62	—
Xin Yi	45.74	—
Chuan Lian Zi	46.70	—
Dang Gui	46.79	—
Yu Zhu	46.84	—
Bai Mao Gen	46.90	—
Han Lian Cao	47.10	—
She Gan	47.21	—
Dang Shen	48.03	—
Bai He	48.49	—
Lu Gen	48.63	—
Zhi Shi	48.73	—

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
62720	62720	0.00	6.00
62719	62719	0.00	6.86
62362	62362	0.00	11.53
62361	62361	0.00	11.66

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 2.1-2020-01

Validated substance/substance group **Qin Jiao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60222-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	25 721
Type B	0	40	0	13 127
Type C	1	0	2	850

The substance/substance group *Qin Jiao* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.8837 % (> 99.2976 %)	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
Jie Geng	5.54	—
Dang Shen	6.40	—
Gou Qi Zi	11.66	—
Huang Qi	12.48	—
Mu Gua	13.12	—
Bai Zhu	13.22	—
San Leng	14.89	—
Dang Gui Wei	15.07	—
She Gan	15.44	—
Long Dan (Cao)	15.98	—
Chai Hu	18.32	—
Sang Zhi	18.83	—
Di Gu Pi	19.07	—
Mu Dan Pi	21.16	—
Bai Shao Yao	21.46	—
Bai Mao Gen	22.03	—
(Huai) Niu Xi	22.06	—
Zhi Mu	22.81	—
Chuan Xiong	23.80	—
Gui Zhi	24.68	—
Cang Zhu	24.88	—
Dang Gui	25.35	—
Chi Shao (Yao)	25.60	—
Du Huo	25.68	—
Zi Wan	26.30	—
Yuan Zhi	27.60	—
Bai He	27.64	—
Gan Jiang	27.83	—
Da Zao	27.99	—
Lai Fu Zi	29.60	—
Mang Xiao	30.00	—
Qing Pi	30.02	—
Ba Ji Tian	30.31	—
Mai Men Dong	30.64	—
Lu Gen	31.99	—
Gua Lou	33.10	—
(Shi) Chang Pu	33.53	—
Tian Hua Fen	33.74	—
Sheng Jiang	33.89	—
Lian Qiao	34.04	—
Fang Feng	34.21	—
San Qi	34.31	—
Sha Shen (Bei)	35.22	—
Chuan Lian Zi	36.99	—
Xiang Fu	37.66	—
E Zhu	39.76	—
Chen Pi	40.16	—
Ji Li	40.55	—
Mai Ya	42.29	—
Ban Zhi Lian	43.46	—
Huo Ma Ren	43.77	—
Cang Er Zi	44.20	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ze Xie	45.34	—
Tu Fu Ling	45.84	—
Ku Shen	46.02	—
Xian He Cao	46.23	—
Zhi Ke	46.25	—
Lian Zi	46.28	—
Shan Yao	46.31	—
Zi Su Zi	46.37	—
Niu Bang Zi	46.40	—
Hou Po	46.40	—
Xie Bai	46.42	—
Hong Jing Tian	46.56	—
Jiang Huang	46.68	—
Wu Yao	48.16	—
Yan Hu Suo	49.23	—
Bai Bian Dou	49.43	—
Gan Cao	49.73	—
Rou Gui	50.43	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/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
62736	62736	0.00	5.91
62735	62735	0.00	5.54

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 2.1-2020-01

Validated substance/substance group **Qing Hao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60198-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G031H0810821	62769	40	from supplier
PhytoComm	Qing Hao	G031H0810821	62770	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G031H0810821	62769 [†]	20
PhytoComm	Qing Hao	G031H0810821	62770 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

The substance/substance group *Qing Hao* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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
Gou Teng	12.14	—
Pu Gong Ying	20.14	—
Xin Yi	27.33	—
Sha Ren	30.91	—
Huang Lian	34.61	—
Huang Bo	35.18	—
Yin Chen Hao	35.85	—
Nü Zhen Zi	36.36	—
Jing Jie	36.56	—
Yin Yang Huo	37.65	—
Xuan Shen	39.31	—
Dan Zhu Ye	41.84	—
Yi Mu Cao	42.91	—
Wu Jia Pi	42.94	—
Xu Duan	43.09	—
Sang Ji Shend	45.31	—
Qiang Huo	45.76	—
Jin Qian Cao	47.59	—
Wu Zhu Yu	47.86	—
(Sheng) Di Huang	48.43	—
Ye Jiao Teng	49.54	—
Hua Shi	49.83	—

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
62769	62769	0.00	12.14
62770	62770	0.00	12.16

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 2.1-2020-01

Validated substance/substance group **Qing Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50371-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Qing Pi	2	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	G074H0809821	62379	40	from supplier
PhytoComm	Qing Pi	G074H0809821	62380	40	from supplier
PhytoComm	Qing Pi	G074HS361SK1	62671	40	from supplier
PhytoComm	Qing Pi	G074HS361SK1	62672	40	from supplier

Validation samples

A total of 39 821 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 *Qing Pi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G074H0809821	62379 [†]	20
PhytoComm	Qing Pi	G074H0809821	62380 [†]	20
PhytoComm	Qing Pi	G074HS361SK1	62671 [†]	20
PhytoComm	Qing Pi	G074HS361SK1	62672 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	g074h100713	1
PhytoComm	Qing Pi	G074H0809621	1

- 851 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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	2	851

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Fang Ji	7.59	—
Huang Bai	8.70	—
Wu Wei Zi	11.52	—
Zhi Ke	12.35	—
Sang Ye	13.38	—
Gua Lou	13.41	—
Sang Zhi	14.39	—
She Gan	14.63	—
Shan Yao	14.74	—
Zi Wan	14.94	—
Niu Bang Zi	15.26	—
Du Zhong	16.06	—
Qiang Huo	16.14	—
Zi Hua Di Ding	16.21	—
Jiao Gu Lan	16.23	—
Ze Xie	16.34	—
Di Gu Pi	16.43	—
Dang Gui Wei	16.78	—
Chai Hu	16.90	—
Mao Dong Qing	18.44	—
Lian Qiao	18.55	—
Gou Qi Zi	20.00	—
Dan Shen	20.38	—
Jin Yin Hua	20.66	—
Hou Po	21.56	—
Xiang Fu	21.82	—
Hong Hua	21.82	—
(Huai) Niu Xi	21.85	—
Huang Lian	21.95	—
Hong Jing Tian	22.04	—
Shan Yu Rou	22.17	—
(Shi) Chang Pu	22.65	—
Gan Jiang	22.72	—
Yi Mu Cao	23.04	—
Wu Mei	23.20	—
Yin Yang Huo	23.30	—
Du Huo	23.30	—
Bai Mao Gen	23.55	—
Yan Hu Suo	23.70	—
Ma Huang	23.91	—
Cang Zhu	23.95	—
Jie Geng	24.63	—
Pi Pa Ye	24.79	—
Bai Zhu	24.95	—
Chuan Xiong	25.18	—
Gui Zhi	25.44	—
Ku Shen	25.67	—
Ye Jiao Teng	26.13	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Bian Dou	26.46	—
Cang Er Zi	27.19	—
Tian Hua Fen	27.26	—
Xi Xian Cao	27.44	—
Mu Dan Pi	27.51	—
Wei Ling Xian	27.75	—
Bai Shao Yao	27.77	—
Huo Ma Ren	27.78	—
Ren Dong Teng	27.89	—
Chi Shao (Yao)	28.62	—
Bai Xian Pi	28.73	—
Huang Qi	28.82	—
San Leng	29.91	—
He Huan Pi	30.17	—
Jiang Huang	30.20	—
Dan Dou Chi	30.21	—
Ji Li	30.63	—
Wu Yao	30.65	—
Zhi Zi	31.29	—
Yu Jin	31.33	—
Xia Ku Cao	31.45	—
Fu Zi	32.29	—
Mang Xiao	32.31	—
Ling Zhi	32.67	—
Pu Gong Ying	32.72	—
Dang Shen	33.04	—
Lu Gen	33.05	—
Tu Fu Ling	33.82	—
Che Qian Zi	34.53	—
Dang Gui	34.91	—
Xie Bai	34.93	—
Qin Jiao	35.07	—
Sha Shen (Bei)	35.35	—
Mu Gua	35.37	—
Suan Zao Ren	35.60	—
Ji Xue Teng	36.03	—
Lian Zi	36.42	—
E Zhu	36.59	—
Fang Feng	36.81	—
Yuan Zhi	37.17	—
Gan Cao	37.63	—
Sheng Jiang	37.65	—
Huang Qin	37.68	—
Zi Su Zi	38.29	—
Zhi Mu	38.73	—
Tao Ren	38.80	—
Lai Fu Zi	38.99	—
(Bai) Dou Kou	39.07	—
San Qi	39.42	—
Sha Ren	40.20	—
Dong Gua Zi	40.71	—
Chuan Lian Zi	40.83	—
Chen Pi	40.98	—
Shen Qu	41.49	—
Zhe Bei Mu	41.78	—
Nü Zhen Zi	42.51	—
Tai Zi Shen	42.61	—
Bai Zi Ren	43.04	—
Xin Yi	43.41	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Dan (Cao)	43.60	—
Gou Teng	43.72	—
Huang Bo	44.69	—
Xian He Cao	45.16	—
Yu Zhu	46.16	—
Ban Xia (Jiang)	46.26	—
Ban Zhi Lian	47.11	—
Ju Hua	47.33	—
Chuang Mu Xiang	48.69	—

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
62379	62379	0.00	13.41
62672	62672	0.00	7.59
62671	62671	0.00	8.70
62380	62380	0.00	14.07

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 2.1-2020-01

Validated substance/substance group **Ren Dong Teng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60104-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	0	79	1	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

The substance/substance group *Ren Dong Teng* 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 % (> 99.9703 %)	98.7500 % (> 95.0000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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	4.86	—
Bai Bian Dou	5.53	—
Shen Qu	6.36	—
Bai Xian Pi	6.94	—
Ye Jiao Teng	8.22	—
Lian Zi	9.20	—
Dan Dou Chi	9.29	—
Ji Li	9.62	—
Yu Jin	10.01	—
Ji Xue Teng	10.12	—
Rou Gui	10.14	—
Huo Ma Ren	10.15	—
Ban Xia (Jiang)	10.62	—
Gou Teng	10.70	—
Yuan Zhi	10.71	—
Chai Hu	10.97	—
Chuan Xiong	11.17	—
Hou Po	11.38	—
Ma Huang Gen	11.42	—
Mao Dong Qing	11.43	—
Shan Yao	11.64	—
Lian Qiao	12.17	—
Pi Pa Ye	12.38	—
He Huan Pi	12.42	—
Yin Yang Huo	12.81	—
Yan Hu Suo	13.09	—
Che Qian Zi	13.14	—
Tu Fu Ling	13.64	—
Dan Shen	13.84	—
Fu Zi	13.94	—
Tian Hua Fen	14.02	—
Fu Xiao Mai	14.03	—
Fu Ling	14.12	—
Bai Shao Yao	14.28	—
Sheng Jiang	14.54	—
Chen Pi	15.10	—
Gui Zhi	15.74	—
Yi Yi Ren	16.77	—
Tao Ren	17.20	—
Tai Zi Shen	17.41	—
Jiao Gu Lan	17.72	—
Gua Lou	18.20	—
Zhe Bei Mu	18.72	—
Bai Zi Ren	18.85	—
Jin Yin Hua	19.15	—
Fo Shou	19.19	—
Ma Huang	19.28	—
Qiang Huo	19.85	—
Hong Jing Tian	20.02	—
Fang Ji	20.05	—
Ci Wu Jia	21.14	—
Di Gu Pi	22.47	—
Zhu Ru	22.75	—
Cang Zhu	23.71	—
Qing Pi	23.90	—
Fu Shen	24.39	—
Wu Wei Zi	25.21	—
Du Zhong	26.04	—
Da Zao	26.06	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tian Men Dong	26.68	—
Mu Dan Pi	27.47	—
Long Yan Rou	27.82	—
Wang Bu Liu Xing	28.06	—
San Qi	28.25	—
Suan Zao Ren	28.43	—
Yu Zhu	29.30	—
Zhi Ke	29.91	—
Huang Lian	29.94	—
Gou Qi Zi	30.60	—
Dang Gui Wei	31.56	—
Zi Hua Di Ding	31.71	—
Dang Gui	34.87	—
Shan Yu Rou	35.46	—
Mang Xiao	35.60	—
Zhi Mu	36.25	—
Dang Shen	36.59	—
Sang Zhi	37.92	—
Huang Qin	37.99	—
Jie Geng	39.68	—
Chuang Mu Xiang	39.96	—
Ze Xie	41.00	—
Chi Shao (Yao)	41.88	—
Huang Bai	42.49	—
Bai Zhu	45.42	—
Bai He	45.48	—
Ban Zhi Lian	46.33	—
Gan Cao	46.65	—
Pu Gong Ying	47.84	—
Cang Er Zi	48.58	—
Chuan Lian Zi	48.80	—
Zi Su Zi	49.09	—

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	6.94
62696	62696	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 2.1-2020-01

Validated substance/substance group	Rou Cong Rong
Substance class	TCM - Granulated herbal extracts (PhytoComm)
Report date	12/03/2020
Report number	60182-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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
Hu Zhang	17.58	—
Bai Hua She She Cao	17.79	—
Han Lian Cao	20.08	—
Zhi Shi	20.30	—
Xin Yi	23.66	—
Sha Ren	23.71	—
Pu Gong Ying	24.04	—
Huang Lian	25.05	—
Sang Ye	25.46	—
Xia Ku Cao	25.79	—
Fu Zi	26.45	—
Yi Mu Cao	26.99	—
Dan Zhu Ye	27.00	—
Gui Zhi	27.10	—
Yin Yang Huo	28.33	—
Gu Sui Bu	29.36	—
Bu Gu Zhi	29.50	—
Gou Teng	29.68	—
Xi Xian Cao	29.80	—
Wei Ling Xian	30.04	—
Huang Bo	31.12	—
Sang Ji Shend	32.17	—
Cang Er Zi	32.60	—
Jiang Huang	32.75	—
Tu Fu Ling	33.04	—
Nü Zhen Zi	33.81	—
Shan Yao	35.80	—
Qiang Huo	35.91	—
Shen Qu	36.24	—
Xuan Shen	36.65	—
Wu Zhu Yu	37.19	—
Xu Duan	37.76	—
Yin Chen Hao	39.08	—
He Huan Pi	39.21	—
Ju Hua	40.08	—
Wu Wei Zi	41.79	—
Qing Pi	42.14	—
Wu Yao	42.49	—
Hong Jing Tian	44.91	—
Ma Bian Cao	45.43	—
Xian He Cao	46.45	—
Dan Shen	46.65	—
Jing Jie	47.04	—
Sang Bai Pi	48.98	—
Wu Mei	49.37	—
(Shi) Chang Pu	49.77	—

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
62668	62668	0.00	18.37
62667	62667	0.00	17.58

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 2.1-2020-01

Validated substance/substance group **Rou Gui**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60169-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G069HS204QM1	62187	40	from supplier
PhytoComm	Rou Gui	G069HS204QM1	62188	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G069HS204QM1	62187 [†]	20
PhytoComm	Rou Gui	G069HS204QM1	62188 [†]	20
PhytoComm	Rou Gui	G069HS204RN1	62271 [†]	20
PhytoComm	Rou Gui	G069HS204RN1	62272 [†]	20
PhytoComm	Rou Gui	G069HS204RM1	62443 [†]	20
PhytoComm	Rou Gui	G069HS204RM1	62444 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	H0617022	1
Phytocomm	Rou Gui	g069h0617123	1
Phytocomm	Rou Gui	G069H0617323	1
PhytoComm	Rou Gui	G069H0617123	2
PhytoComm	Rou Gui	G069H0617323	1

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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	236	4	25 552
Type B	7	104	16	13 040
Type C	1	0	6	846

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	99.9709 % (> 99.9560 %)	98.3333 % (> 97.0833 %)
Type B	99.9273 % (> 99.8984 %)	86.6667 % (> 84.1667 %)
Type C	99.7674 % (> 99.1792 %)	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
Lian Zi	2.90	—
Fu Xiao Mai	3.36	—
Zhu Ru	3.39	—
Ban Xia (Jiang)	5.71	—
Ji Li	6.57	—
Yi Yi Ren	6.83	—
Gui Zhi	7.10	—
Bai Bian Dou	7.42	—
Sheng Jiang	7.57	—
Ma Huang Gen	7.75	—
He Huan Pi	8.25	—
Bai Shao Yao	8.65	—
Gou Teng	8.94	—
Di Gu Pi	9.44	—
Bai Zi Ren	10.12	—
Fu Ling	11.01	—
Fu Shen	11.05	—
Ji Xue Teng	11.18	—
Ci Wu Jia	11.24	—
Ling Zhi	11.51	—
Shen Qu	11.83	—
Wang Bu Liu Xing	11.86	—
Tian Hua Fen	12.02	—
Zhe Bei Mu	12.68	—
Tao Ren	12.69	—
Tai Zi Shen	13.46	—
Huo Ma Ren	13.68	—
Fo Shou	14.28	—
Shan Yao	14.70	—
Gua Lou	15.07	—
Bai Xian Pi	15.52	—
Ye Jiao Teng	16.31	—
Pi Pa Ye	16.39	—
Lian Qiao	17.16	—
Yuan Zhi	18.32	—
Ren Dong Teng	18.43	—
Tian Men Dong	18.89	—
Long Yan Rou	20.21	—
Yu Jin	20.36	—
Chen Pi	20.43	—
Da Zao	21.93	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Zi	22.24	—
Chuan Xiong	22.64	—
Yan Hu Suo	23.58	—
Tu Fu Ling	24.50	—
Dang Shen	25.73	—
Dan Shen	25.83	—
Shan Yu Rou	26.73	—
Dang Gui Wei	26.97	—
Mao Dong Qing	27.07	—
San Qi	27.34	—
Wu Wei Zi	28.09	—
Fang Ji	28.41	—
Jiao Gu Lan	28.49	—
Che Qian Zi	28.58	—
Jin Yin Hua	29.28	—
Dan Dou Chi	29.88	—
Hou Po	30.36	—
Zhi Mu	30.51	—
Yin Yang Huo	31.65	—
Chai Hu	31.96	—
Yu Zhu	32.21	—
Cang Zhu	32.44	—
Dang Gui	32.51	—
Zhi Ke	33.66	—
Huang Qin	33.99	—
Ma Huang	34.20	—
Mang Xiao	37.42	—
Suan Zao Ren	38.39	—
Hong Jing Tian	38.39	—
Gou Qi Zi	41.89	—
Qing Pi	42.46	—
Qiang Huo	42.66	—
Huang Lian	43.62	—
Bai Zhu	44.95	—
Du Zhong	45.12	—
Ze Xie	45.49	—
Sang Zhi	46.14	—
Jie Geng	46.70	—
Mu Dan Pi	47.34	—
Gan Cao	47.87	—

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
62187	62187	0.00	3.39
62271	62271	0.00	3.55
62443	62443	0.00	3.78
62272	62272	0.00	2.90
62444	62444	0.00	5.73
62188	62188	0.00	3.36

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 2.1-2020-01

Validated substance/substance group **Ru Xiang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50370-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Ru Xiang; Olibanum

Special notes

When selecting the *Ru Xiang* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ru Xiang	1	0	2

Second-stage model

For differentiation of the substance/substance group *Ru 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 *Ru Xiang*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Ru Xiang	G174H0820621	61935	40	from supplier
PhytoComm	Ru Xiang	G174H0820621	61936	40	from supplier

Validation samples

A total of 39 821 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 *Ru Xiang*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Ru 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	Ru Xiang	G174H0820621	61935 [†]	20
PhytoComm	Ru Xiang	G174H0820621	61936 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Ru 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	Ru Xiang	G174H0820231	1
Phytocomm	Ru Xiang	G174H0820421	1
PhytoComm	Ru Xiang	G174H0820421	2

- 849 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 *Ru 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 *Ru 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	25 721
Type B	0	40	0	13 127
Type C	0	0	4	849

The substance/substance group *Ru 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8246 %)	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 *Ru 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
Hua Shi	65.66	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Ru 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
61935	61935	0.00	65.73
61936	61936	0.00	65.66

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 2.1-2020-01

Validated substance/substance group **San Leng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60029-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

San Leng; Sparganii rhizoma

Special notes

When selecting the *San Leng* substance/substance group, the following information is displayed to the user:

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
San Leng	1	0	1

Second-stage model

For differentiation of the substance/substance group *San Leng* 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 Leng*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	San Leng	G270H0311721	62235	40	from supplier
PhytoComm	San Leng	G270H0311721	62236	40	from supplier

Validation samples

A total of 39 821 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 Leng*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Leng*.
- 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 Leng	G270H0311721	62235 [†]	20
PhytoComm	San Leng	G270H0311721	62236 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *San Leng*.
- 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 Leng	g270h0311221	1
PhytoComm	San Leng	G270H0311221	1

- 851 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 *San Leng* 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 Leng* 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	25 721
Type B	0	40	0	13 127
Type C	4	1	1	847

The substance/substance group *San Leng* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.7132 % (> 99.1271 %)	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 Leng* 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
Di Gu Pi	10.37	—
Sang Zhi	10.61	—
Gui Zhi	11.39	—
Huang Qi	14.96	—
Jie Geng	15.52	—
Bai Zhu	16.82	—
Mu Dan Pi	17.02	—
Lu Gen	17.02	—
She Gan	17.13	—
Du Huo	18.52	—
Tian Hua Fen	18.70	—
Long Dan (Cao)	19.08	—
Gou Qi Zi	19.12	—
Dang Gui Wei	20.59	—
Dang Gui	21.26	—
Bai Shao Yao	21.41	—
Dang Shen	21.51	—
Qin Jiao	22.07	—
Bai Mao Gen	22.32	—
Cang Zhu	22.67	—
Chuan Xiong	23.75	—
Ban Zhi Lian	23.80	—
(Shi) Chang Pu	24.42	—
Chi Shao (Yao)	24.83	—
Sheng Jiang	25.32	—
Chai Hu	25.61	—
Xiang Fu	25.62	—
(Huai) Niu Xi	26.01	—
Gan Jiang	26.34	—
Bai He	26.36	—
Gua Lou	26.64	—
Chuan Lian Zi	27.77	—
Qing Pi	28.72	—
Sha Shen (Bei)	28.94	—
Mai Men Dong	30.26	—
Mu Gua	31.02	—
Zhi Mu	31.07	—
Mang Xiao	31.08	—
San Qi	31.56	—
Ba Ji Tian	31.66	—
Zi Wan	32.03	—
Ji Li	34.58	—
E Zhu	35.21	—
Tu Fu Ling	35.71	—
Lian Qiao	35.79	—
Lai Fu Zi	35.82	—
Mai Ya	35.94	—
Huo Ma Ren	37.06	—
Shan Yao	37.81	—
Da Zao	37.87	—
Wu Yao	38.32	—
Zhi Ke	38.78	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	38.87	—
Fang Feng	39.22	—
Lian Zi	39.31	—
Yi Mu Cao	40.04	—
Ze Xie	40.15	—
Bai Xian Pi	40.30	—
Cang Er Zi	40.54	—
Tian Men Dong	40.85	—
Ji Xue Teng	41.00	—
Huang Lian	41.18	—
Rou Gui	41.67	—
Jiang Huang	42.19	—
Wu Wei Zi	42.38	—
Jin Yin Hua	43.09	—
Gou Teng	43.36	—
Zi Su Zi	44.07	—
Chen Pi	44.16	—
Hong Jing Tian	44.54	—
Zhu Ling	44.64	—
Yan Hu Suo	45.30	—
Hou Po	45.66	—
Xie Bai	45.76	—
Fu Ling	45.76	—
Dan Shen	45.91	—
Zhe Bei Mu	46.06	—
Ye Jiao Teng	46.11	—
Gan Cao	46.30	—
Ku Shen	46.45	—
Fo Shou	46.96	—
Dan Dou Chi	47.37	—
Ren Dong Teng	47.50	—
Niu Bang Zi	48.00	—
Xian He Cao	48.54	—
Shan Yu Rou	49.21	—
Bai Bian Dou	49.92	—
Sha Ren	49.92	—
Shen Qu	50.06	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *San Leng* 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
62236	62236	0.00	10.88
62235	62235	0.00	10.37

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 2.1-2020-01

Validated substance/substance group **San Qi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60415-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
San Qi	3	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	G203H0312621	62135	40	from supplier
PhytoComm	San Qi	G203H0312621	62136	40	from supplier
PhytoComm	San Qi	G203H0312721	62245	40	from supplier
PhytoComm	San Qi	G203H0312721	62246	40	from supplier
PhytoComm	San Qi	G203HS030SL1	62701	40	from supplier
PhytoComm	San Qi	G203HS030SL1	62702	40	from supplier

Validation samples

A total of 39 821 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 *San Qi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G203H0312621	62135 [†]	20
PhytoComm	San Qi	G203H0312621	62136 [†]	20
PhytoComm	San Qi	G203H0312721	62245 [†]	20
PhytoComm	San Qi	G203H0312721	62246 [†]	20
PhytoComm	San Qi	G203HS030SL1	62701 [†]	20
PhytoComm	San Qi	G203HS030SL1	62702 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	2	0	1	850

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 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	99.6512 % (> 99.0683 %)	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 *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
Chuan Xiong	7.96	—
Lian Qiao	8.68	—
Shan Yao	8.89	—
Yuan Zhi	9.86	—
Dang Shen	9.86	—
Pi Pa Ye	10.20	—
Bai Bian Dou	10.50	—
Bai Shao Yao	10.58	—
Shan Yu Rou	11.03	—
Cang Zhu	11.55	—
Bai Zi Ren	11.78	—
Ji Li	12.36	—
Chuan Lian Zi	12.43	—
Sheng Jiang	12.97	—
Suan Zao Ren	14.92	—
Lai Fu Zi	15.66	—
Huo Ma Ren	16.01	—
Jiao Gu Lan	16.03	—
Ban Xia (Jiang)	16.16	—
Zhe Bei Mu	16.23	—
Zhi Mu	16.34	—
Jin Yin Hua	16.38	—
Dang Gui	16.56	—
Gua Lou	16.81	—
Jie Geng	17.00	—
Gou Qi Zi	17.52	—
Lian Zi	17.58	—
Dang Gui Wei	17.68	—
Bai He	18.32	—
Gan Jiang	18.34	—
Tian Hua Fen	18.39	—
Di Gu Pi	18.47	—
Shen Qu	18.49	—
Ren Dong Teng	18.50	—
Mai Men Dong	18.80	—
Chen Pi	19.25	—
Fu Xiao Mai	19.39	—
Bai Zhu	19.46	—
Tai Zi Shen	19.56	—
Yu Zhu	19.69	—
Mao Dong Qing	19.97	—
Rou Gui	20.05	—
He Huan Pi	20.30	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Zhi	20.53	—
Wu Wei Zi	20.68	—
Mu Dan Pi	20.69	—
Hou Po	21.27	—
Qing Pi	21.28	—
Mu Gua	21.45	—
Ye Jiao Teng	21.65	—
Zi Su Zi	21.70	—
Chuang Mu Xiang	21.83	—
Zhu Ru	21.87	—
Fu Zi	21.95	—
Gui Zhi	22.58	—
Cang Er Zi	23.07	—
Ling Zhi	23.68	—
Fu Ling	23.93	—
E Zhu	24.79	—
Ci Wu Jia	25.21	—
Fang Ji	25.30	—
Dan Shen	25.75	—
Zhi Ke	26.01	—
Tu Fu Ling	26.50	—
Mang Xiao	27.09	—
Tao Ren	27.21	—
Gou Teng	27.59	—
Ji Xue Teng	27.92	—
Xiang Fu	27.94	—
Yi Yi Ren	28.08	—
Mo Yao	28.28	—
Da Zao	28.67	—
She Gan	29.76	—
Yan Hu Suo	29.99	—
Long Dan (Cao)	30.07	—
Huang Qin	30.66	—
Bai Xian Pi	30.94	—
Fo Shou	31.35	—
Qiang Huo	31.58	—
(Shi) Chang Pu	31.63	—
Gan Cao	31.65	—
Chai Hu	31.70	—
Tian Men Dong	31.86	—
San Leng	31.90	—
Che Qian Zi	32.24	—
Mai Ya	32.31	—
Yin Yang Huo	32.72	—
Ku Shen	33.04	—
Wang Bu Liu Xing	33.84	—
Hong Jing Tian	33.96	—
Long Yan Rou	34.02	—
Ma Huang	35.13	—
Dan Dou Chi	35.43	—
Yu Jin	35.49	—
Huang Qi	35.62	—
Qin Jiao	36.29	—
Jiang Huang	36.37	—
Niu Bang Zi	36.42	—
Du Huo	36.83	—
Xie Bai	38.32	—
Zi Wan	38.71	—
Ban Zhi Lian	39.44	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ba Ji Tian	39.54	—
Bai Mao Gen	40.00	—
Du Zhong	40.64	—
Ma Huang Gen	40.66	—
Dong Gua Zi	41.62	—
Ze Xie	41.63	—
Zi Hua Di Ding	42.21	—
Chi Shao (Yao)	42.52	—
Lu Gen	43.21	—
Yi Mu Cao	43.29	—
Huang Lian	45.09	—
Fang Feng	45.75	—
(Huai) Niu Xi	46.41	—
Sha Shen (Bei)	46.65	—
Wei Ling Xian	47.02	—
Fu Shen	48.26	—
(Bai) Dou Kou	49.45	—

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
62245	62245	0.00	12.99
62135	62135	0.00	13.05
62702	62702	0.00	7.96
62701	62701	0.00	8.88
62136	62136	0.00	12.43
62246	62246	0.00	12.97

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 2.1-2020-01

Validated substance/substance group **Sang Bai Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50287-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G161H1042822	62471	40	from supplier
PhytoComm	Sang Bai Pi	G161H1042822	62472	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G161H1042822	62471 [†]	20
PhytoComm	Sang Bai Pi	G161H1042822	62472 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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

- 849 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 *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	25 721
Type B	0	40	0	13 127
Type C	1	0	4	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.7674 % (> 99.1797 %)	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
Dan Zhu Ye	14.49	—
Nü Zhen Zi	16.41	—
Sang Ji Shend	17.34	—
Jing Jie	18.02	—
Bu Gu Zhi	19.23	—
Xu Duan	21.70	—
He Shou Wu	22.40	—
Gu Sui Bu	22.75	—
Ma Bian Cao	25.64	—
Yin Chen Hao	25.85	—
Yin Yang Huo	26.57	—
He Huan Pi	26.85	—
Jin Qian Cao	27.21	—
Xian He Cao	27.49	—
Du Zhong	28.20	—
Yi Mu Cao	28.87	—
(Sheng) Di Huang	30.89	—
Ge Gen	31.79	—
Wu Jia Pi	33.29	—
Chen Pi	36.30	—
Gui Zhi	36.54	—
Yu Jin	38.01	—
Fang Feng	38.65	—
Gou Teng	38.92	—
Qiang Huo	39.43	—
Huang Bo	40.06	—
Sha Ren	40.10	—
Mang Xiao	40.22	—
Dan Shen	40.50	—
Fu Zi	40.86	—
Hong Jing Tian	41.63	—
Wu Zhu Yu	41.95	—
Hou Po	42.25	—
Pu Gong Ying	42.93	—
Lu Gen	43.78	—
Hu Zhang	43.83	—
Ji Li	44.17	—
Wu Yao	45.75	—
Jiang Huang	46.45	—
Huang Lian	47.32	—
Hua Shi	49.27	—

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
62472	62472	0.00	14.49
62471	62471	0.00	16.32

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 2.1-2020-01

Validated substance/substance group **Sang Ji Shend**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60142-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Sang Ji Shend	2	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

Validation samples

A total of 39 821 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 Ji Shend*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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

- 13 087 spectra from a total of 318 batches from further 172 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	G146H1045522	4
PhytoComm	Sang Ji Shend	G146H1045322	1

- 847 spectra from 12 *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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	2	4	847

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8235 %)	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 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
Sang Bai Pi	15.88	—
Gu Sui Bu	18.82	—
Ge Gen	23.41	—
Xian He Cao	24.12	—
Hong Jing Tian	25.92	—
Xu Duan	25.98	—
He Shou Wu	26.82	—
Wu Zhu Yu	31.78	—
Yin Yang Huo	32.12	—
Zhi Shi	32.68	—
Nü Zhen Zi	33.08	—
He Huan Pi	34.32	—
Yu Jin	34.72	—
Yi Mu Cao	35.04	—
Dan Zhu Ye	36.14	—
Bu Gu Zhi	36.40	—
Mang Xiao	37.03	—
Yin Chen Hao	37.31	—
Dan Shen	38.52	—
Jing Jie	38.77	—
Chen Pi	38.82	—
Ma Bian Cao	39.50	—
Qiang Huo	39.50	—
Hou Po	40.86	—
Fang Feng	42.07	—
Lu Gen	42.83	—
Gui Zhi	43.92	—
Fu Zi	44.60	—
Huang Bo	44.96	—
Huang Jing	45.31	—
Sha Ren	45.33	—
Wu Yao	45.51	—
Hua Shi	49.26	—
Ji Li	50.02	—

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	17.37
62543	62543	0.00	15.88
62544	62544	0.00	16.12
62346	62346	0.00	17.48

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 2.1-2020-01

Validated substance/substance group **Sang Ye**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60079-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Sang Ye	1	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

Validation samples

A total of 39 821 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 Ye*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 848 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 *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	80	0	25 721
Type B	1	40	0	13 126
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	99.9964 % (> 99.9676 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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 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
Pu Gong Ying	11.98	—
Shan Yao	13.12	—
Huang Lian	13.25	—
Xia Ku Cao	17.66	—
Xin Yi	18.01	—
Tu Fu Ling	18.60	—
Xi Xian Cao	18.92	—
Wu Mei	19.44	—
Wu Wei Zi	20.04	—
Qiang Huo	20.77	—
Nü Zhen Zi	20.90	—
Gui Zhi	21.05	—
Qing Pi	21.33	—
Fu Zi	21.97	—
Dan Shen	21.99	—
Du Zhong	22.41	—
Cang Er Zi	23.28	—
Mao Dong Qing	24.51	—
Zhi Ke	24.85	—
Fang Ji	25.07	—
Jiang Huang	26.94	—
Hong Jing Tian	27.75	—
Hu Zhang	28.25	—
Huang Bai	28.79	—
Hong Hua	29.04	—
Yin Yang Huo	30.24	—
Hou Po	30.46	—
Ye Jiao Teng	31.46	—
Yi Mu Cao	32.20	—
Jiao Gu Lan	32.25	—
Ma Huang	32.63	—
Han Lian Cao	33.73	—
Yu Jin	33.75	—
Chai Hu	33.83	—
Shen Qu	34.99	—
Jin Yin Hua	36.82	—
Ze Xie	37.08	—
Bai Xian Pi	37.67	—
Shan Yu Rou	37.79	—
Pi Pa Ye	37.82	—
Di Gu Pi	38.42	—
Zi Hua Di Ding	38.87	—
Wu Yao	39.75	—
Yan Hu Suo	40.00	—
Chuan Xiong	40.68	—
Xu Duan	40.81	—
Wang Bu Liu Xing	40.83	—
Bu Gu Zhi	41.24	—
Bai Shao Yao	41.41	—
Mang Xiao	42.06	—
Bai Zhu	42.16	—
He Huan Pi	42.31	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Dou Chi	42.61	—
Xiang Fu	42.80	—
Sha Ren	42.90	—
Bai Mao Gen	43.38	—
Chi Shao (Yao)	43.73	—
Bai Hua She She Cao	44.31	—
Lian Qiao	44.72	—
Che Qian Zi	44.84	—
Ren Dong Teng	45.73	—
Gan Cao	45.82	—
Ling Zhi	46.06	—
Gua Lou	46.29	—
(Shi) Chang Pu	46.82	—
Gou Qi Zi	47.69	—
Wei Ling Xian	47.77	—
Lu Gen	47.77	—
Lian Zi	48.21	—
Ji Xue Teng	48.28	—
Mu Dan Pi	48.39	—
E Zhu	48.70	—
Bai Bian Dou	49.17	—
Sang Zhi	49.26	—
Huang Qin	49.62	—
Zhi Zi	49.66	—

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	13.12
62462	62462	0.00	11.98

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 2.1-2020-01

Validated substance/substance group **Sang Zhi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60091-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Sang Zhi	3	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	G164H1044621	61969	40	from supplier
PhytoComm	Sang Zhi	G164H1044621	61970	40	from supplier
PhytoComm	Sang Zhi	G164H1044721	62359	40	from supplier
PhytoComm	Sang Zhi	G164H1044721	62360	40	from supplier
PhytoComm	Sang Zhi	G164H1044821	62473	40	from supplier
PhytoComm	Sang Zhi	G164H1044821	62474	40	from supplier

Validation samples

A total of 39 821 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 Zhi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 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	G164H1044621	61969 [†]	20
PhytoComm	Sang Zhi	G164H1044621	61970 [†]	20
PhytoComm	Sang Zhi	G164H1044721	62359 [†]	20
PhytoComm	Sang Zhi	G164H1044721	62360 [†]	20
PhytoComm	Sang Zhi	G164H1044821	62473 [†]	20
PhytoComm	Sang Zhi	G164H1044821	62474 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	240	0	25 561
Type B	0	119	1	13 047
Type C	8	0	0	845

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	99.1667 % (> 96.6667 %)
Type C	98.6728 % (> 98.0890 %)	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
Di Gu Pi	5.81	—
Gui Zhi	6.85	—
Ji Li	9.48	—
Jie Geng	9.70	—
Chuan Lian Zi	10.90	—
Tian Hua Fen	11.06	—
Xiang Fu	11.23	—
San Leng	11.49	—
Gua Lou	12.91	—
Sha Shen (Bei)	13.12	—
Long Dan (Cao)	13.65	—
Qing Pi	13.68	—
E Zhu	14.13	—
Chuan Xiong	14.58	—
Ba Ji Tian	14.70	—
Chai Hu	15.06	—
Dang Gui Wei	15.57	—
(Shi) Chang Pu	15.57	—
Gan Jiang	15.88	—
Dang Gui	16.16	—
Bai Mao Gen	16.18	—
Bai Zhu	16.38	—
Sheng Jiang	17.40	—
She Gan	17.42	—
(Huai) Niu Xi	18.05	—
Du Huo	18.52	—
Lu Gen	18.63	—
Cang Er Zi	18.79	—
Gou Qi Zi	19.13	—
Mu Dan Pi	19.74	—
Huang Qi	20.16	—
Cang Zhu	20.23	—
Dang Shen	20.70	—
Zhi Mu	21.16	—
Huo Ma Ren	21.36	—
Yan Hu Suo	21.75	—
Jiang Huang	21.78	—
Yi Mu Cao	22.07	—
Bai He	22.55	—
Zi Wan	22.71	—
Zi Su Zi	22.82	—
Qin Jiao	22.91	—
Xie Bai	22.93	—
Lai Fu Zi	23.28	—
Shan Yao	23.60	—
Ze Xie	23.90	—
San Qi	24.36	—
Niu Bang Zi	24.88	—
Mang Xiao	24.95	—
Fang Feng	25.67	—
Ban Zhi Lian	25.83	—
Lian Qiao	27.19	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	27.31	—
(Bai) Dou Kou	28.15	—
Mu Gua	28.69	—
Chi Shao (Yao)	29.00	—
Ku Shen	29.12	—
Mai Ya	29.20	—
Yuan Zhi	29.44	—
Hou Po	29.69	—
Chen Pi	30.02	—
He Huan Pi	30.20	—
Sha Ren	30.33	—
Wu Yao	30.37	—
Mai Men Dong	30.44	—
Bai Shao Yao	30.93	—
Wu Wei Zi	32.15	—
Lian Zi	33.13	—
Pu Gong Ying	33.28	—
Xian He Cao	33.40	—
Zhu Ling	33.58	—
Jin Yin Hua	33.63	—
Bai Bian Dou	34.24	—
Huang Lian	34.48	—
Hong Jing Tian	35.39	—
Bai Xian Pi	36.16	—
Zhi Ke	36.35	—
Qiang Huo	37.34	—
Fu Ling	37.77	—
Gou Teng	38.49	—
Tian Men Dong	38.55	—
Rou Gui	38.80	—
Ren Dong Teng	38.84	—
Wang Bu Liu Xing	39.49	—
Da Zao	39.77	—
Wei Ling Xian	40.25	—
Ji Xue Teng	40.33	—
Fang Ji	41.38	—
Yu Jin	41.54	—
Gan Cao	41.98	—
Dong Gua Zi	42.04	—
Fo Shou	42.05	—
Fu Zi	42.19	—
Shan Yu Rou	42.42	—
Mo Yao	42.67	—
Ban Xia (Jiang)	42.87	—
Zhe Bei Mu	42.95	—
Ju Hua	43.37	—
Ye Jiao Teng	43.61	—
Che Qian Zi	44.36	—
Dan Dou Chi	44.37	—
Xin Yi	44.53	—
Zhi Zi	46.02	—
Shen Qu	47.00	—
Dan Shen	47.11	—
Yin Yang Huo	47.68	—
Zi Hua Di Ding	48.00	—
Du Zhong	48.20	—
Yi Yi Ren	49.04	—
Jiao Gu Lan	49.27	—
Pi Pa Ye	49.76	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhu Ru	49.81	–

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
61969	61969	0.00	9.48
62359	62359	0.00	9.44
62473	62473	0.00	5.81
61970	61970	0.00	9.81
62360	62360	0.00	9.10
62474	62474	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 2.1-2020-01

Validated substance/substance group **Sha Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60184-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G018H0939622	62095	40	from supplier
PhytoComm	Sha Ren	G018H0939622	62096	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G018H0939622	62095 [†]	20
PhytoComm	Sha Ren	G018H0939622	62096 [†]	20
PhytoComm	Sha Ren	G018H0939821	62401 [†]	20
PhytoComm	Sha Ren	G018H0939821	62404 [†]	20
PhytoComm	Sha Ren	G018H0939922	62763 [†]	20
PhytoComm	Sha Ren	G018H0939922	62764 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 848 spectra from 12 *Apo-Ident* customers from a total of 511 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	25 561
Type B	0	120	0	13 047
Type C	1	0	5	847

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	99.8837 % (> 99.2957 %)	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
Jing Jie	11.96	—
Dan Zhu Ye	12.14	—
Gui Zhi	13.30	—
Gu Sui Bu	15.04	—
Wu Jia Pi	15.13	—
Fu Zi	15.38	—
Du Zhong	15.99	—
Bu Gu Zhi	16.10	—
Yin Chen Hao	17.66	—
Yu Jin	17.72	—
Chen Pi	18.25	—
Yi Mu Cao	18.25	—
Sang Bai Pi	18.62	—
Wu Yao	18.68	—
Jiang Huang	18.95	—
(Sheng) Di Huang	19.44	—
Wu Zhu Yu	20.11	—
Xu Duan	20.59	—
He Shou Wu	21.07	—
Yu Xing Cao	21.98	—
(Bai) Dou Kou	23.32	—
Xian He Cao	23.47	—
Yin Yang Huo	23.76	—
Sang Ji Shend	24.21	—
Ma Bian Cao	24.54	—
Ji Li	24.73	—
He Huan Pi	26.08	—
Lu Gen	26.13	—
Tu Fu Ling	26.17	—
Xuan Shen	27.63	—
Jin Qian Cao	27.78	—
Hong Jing Tian	28.09	—
Dan Shen	28.96	—
Chai Hu	29.57	—
Hou Po	30.42	—
Mang Xiao	30.44	—
Xiang Fu	30.44	—
Pu Gong Ying	30.93	—
Mu Dan Pi	32.37	—
Fang Feng	32.43	—
(Shi) Chang Pu	32.69	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yan Hu Suo	32.72	—
Qiang Huo	34.12	—
Shan Yao	34.54	—
Sang Ye	35.09	—
Huang Bo	36.09	—
She Gan	36.29	—
Nü Zhen Zi	36.31	—
Sang Zhi	36.48	—
Lian Qiao	36.83	—
E Zhu	37.14	—
Gou Qi Zi	37.41	—
Gua Lou	37.41	—
Huo Ma Ren	38.47	—
Gan Jiang	38.78	—
Hu Zhang	39.45	—
Zi Wan	39.69	—
Zhi Shi	40.11	—
Gou Teng	40.16	—
Xia Ku Cao	40.18	—
Ze Xie	40.18	—
Qing Pi	40.32	—
Ge Gen	40.82	—
Tian Hua Fen	41.70	—
Guang Huo Xiang	42.04	—
Xie Bai	42.10	—
Bai Hua She She Cao	42.16	—
Qing Hao	42.31	—
Chuan Xiong	42.62	—
Huang Jing	43.21	—
Zhi Ke	43.23	—
Wang Bu Liu Xing	43.33	—
Sha Shen (Bei)	43.39	—
Cang Er Zi	43.46	—
Xin Yi	44.12	—
Dang Gui	44.38	—
Huang Lian	45.06	—
Di Gu Pi	45.31	—
Zi Su Zi	45.54	—
Dang Gui Wei	46.16	—
Hua Shi	47.30	—
Ku Shen	47.49	—
Niu Bang Zi	48.23	—
Xi Xian Cao	48.50	—
Shu Di (Huang)	48.92	—
Tu Si Zi	49.00	—
Ba Ji Tian	49.23	—
Han Lian Cao	49.34	—
Wei Ling Xian	49.77	—

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
62095	62095	0.00	12.49
62764	62764	0.00	12.14
62401	62401	0.00	13.32
62404	62404	0.00	13.30
62763	62763	0.00	12.55
62096	62096	0.00	11.96

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 2.1-2020-01

Validated substance/substance group **Sha Shen (Bei)**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60200-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Sha Shen (Bei)	1	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

Validation samples

A total of 39 821 spectra were provided for validation. The results were evaluated separately according to the following sample categories:

Type A All calibration spectra.

- 40 spectra of 1 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 1 different batches.
- 25 761 spectra from a total of 318 batches from further 172 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.

- 79 spectra of 2 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

- 13 088 spectra from a total of 319 batches from further 172 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

- 849 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 *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	40	0	25 761
Type B	0	79	0	13 088
Type C	0	0	4	849

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 % (> 99.9705 %)	100.0000 % (> 85.0000 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.4051 %)
Type C	100.0000 % (> 98.8246 %)	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
E Zhu	9.23	—
Ji Li	10.58	—
Sang Zhi	11.70	—
Yan Hu Suo	13.44	—
Jiang Huang	17.09	—
Sheng Jiang	17.31	—
(Bai) Dou Kou	17.88	—
Di Gu Pi	18.25	—
Chuan Lian Zi	18.51	—
Long Dan (Cao)	18.76	—
Ba Ji Tian	19.77	—
Cang Er Zi	20.03	—
Gui Zhi	21.61	—
Chuan Xiong	21.68	—
San Qi	22.58	—
Mang Xiao	22.82	—
Sha Ren	23.24	—
Yi Mu Cao	23.29	—
Gan Jiang	23.41	—
(Huai) Niu Xi	23.45	—
Chai Hu	23.87	—
Zi Su Zi	24.03	—
Bai He	24.69	—
He Huan Pi	25.08	—
Tu Fu Ling	25.18	—
Wang Bu Liu Xing	25.40	—
Mai Ya	26.29	—
Lai Fu Zi	26.29	—
Zhu Ling	26.61	—
Mu Dan Pi	26.75	—
Lu Gen	27.11	—
Huo Ma Ren	27.13	—
Shan Yao	27.21	—
Dang Gui	27.26	—
She Gan	27.43	—
Xiang Fu	27.50	—
Bai Bian Dou	28.46	—
Du Huo	29.08	—
Tian Hua Fen	29.67	—
Dang Shen	29.71	—
Xie Bai	29.95	—
Gua Lou	30.27	—
Hou Po	30.61	—
(Shi) Chang Pu	30.64	—
Wu Yao	31.19	—
San Leng	31.21	—
Zhi Mu	31.97	—
Chen Pi	33.01	—
Huang Lian	33.05	—
Niu Bang Zi	33.26	—
Yuan Zhi	33.51	—
Dang Gui Wei	33.58	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mo Yao	33.60	—
Jie Geng	33.71	—
Fang Feng	33.92	—
Gou Qi Zi	34.46	—
Qing Pi	34.60	—
Zi Wan	34.66	—
Xian He Cao	35.02	—
Fu Ling	35.43	—
Bai Zhu	35.71	—
Yu Jin	36.11	—
Ban Zhi Lian	37.08	—
Fang Ji	37.51	—
Cang Zhu	37.64	—
Mai Men Dong	37.67	—
Ban Xia (Jiang)	38.29	—
Lian Qiao	38.61	—
Rou Gui	38.69	—
Lian Zi	38.83	—
Bai Mao Gen	40.08	—
Fu Zi	40.28	—
Huang Qi	40.67	—
Xin Yi	41.02	—
Qin Jiao	41.14	—
Bai Xian Pi	42.10	—
Hong Jing Tian	42.89	—
Zhi Ke	44.24	—
Mu Gua	44.34	—
Qiang Huo	44.69	—
Wu Wei Zi	44.84	—
Pu Gong Ying	45.14	—
Ji Xue Teng	45.53	—
Da Zao	45.56	—
Du Zhong	46.99	—
Ze Xie	47.12	—
Gou Teng	47.17	—
Yin Yang Huo	47.38	—
Jin Yin Hua	48.39	—
Dan Dou Chi	48.72	—
Sang Ji Shend	48.90	—
Ku Shen	49.02	—
Tian Men Dong	49.13	—
Sang Bai Pi	50.30	—

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	9.23

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 2.1-2020-01

Validated substance/substance group **Shan Yao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60207-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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

- 13 047 spectra from a total of 317 batches from further 172 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

- 849 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 *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	25 558
Type B	10	118	2	13 037
Type C	3	3	1	846

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	99.9830 % (> 99.9681 %)	99.5833 % (> 98.3333 %)
Type B	99.8837 % (> 99.8548 %)	98.3333 % (> 95.8333 %)
Type C	99.6512 % (> 99.0635 %)	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	3.83	—
Fu Zi	4.63	—
Lian Qiao	5.02	—
Bai Bian Dou	6.07	—
Ren Dong Teng	7.64	—
Ji Li	8.48	—
Shen Qu	9.53	—
Yuan Zhi	10.02	—
Wu Wei Zi	10.06	—
Huo Ma Ren	10.19	—
Bai Zi Ren	10.21	—
Chuan Xiong	10.29	—
Gua Lou	10.70	—
Jiao Gu Lan	12.09	—
Jin Yin Hua	12.29	—
Hou Po	12.29	—
Lian Zi	12.31	—
Tian Hua Fen	12.63	—
Bai Shao Yao	13.17	—
Tai Zi Shen	14.05	—
Zhe Bei Mu	14.46	—
Ban Xia (Jiang)	14.49	—
He Huan Pi	14.65	—
Cang Zhu	15.02	—
Ling Zhi	15.03	—
Dan Shen	15.09	—
Tu Fu Ling	15.40	—
San Qi	15.54	—
Chen Pi	16.11	—
Yan Hu Suo	16.30	—
Mao Dong Qing	16.86	—
Ji Xue Teng	17.03	—
Chai Hu	17.16	—
Fu Ling	17.81	—
Fang Ji	18.03	—
Fu Xiao Mai	18.35	—
Tao Ren	18.47	—
Shan Yu Rou	18.67	—
Che Qian Zi	18.82	—
Cang Er Zi	18.83	—
Sheng Jiang	18.88	—
Di Gu Pi	19.05	—
Bai Xian Pi	19.10	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yin Yang Huo	19.26	—
Rou Gui	19.48	—
Gou Teng	19.64	—
Ye Jiao Teng	20.19	—
Yu Jin	20.36	—
Mu Dan Pi	20.46	—
Dan Dou Chi	20.52	—
Gui Zhi	20.99	—
Hong Jing Tian	21.30	—
Ci Wu Jia	21.59	—
Yu Zhu	21.73	—
Zhu Ru	21.97	—
Ma Huang	22.46	—
Suan Zao Ren	22.71	—
Dang Gui	22.80	—
Gou Qi Zi	22.82	—
Qing Pi	23.06	—
Dang Shen	23.32	—
Dang Gui Wei	23.87	—
Hong Hua	24.22	—
Yi Yi Ren	24.79	—
Qiang Huo	25.44	—
Da Zao	26.06	—
Zhi Ke	26.39	—
Sang Ye	27.12	—
Bai Zhu	28.59	—
Ma Huang Gen	28.67	—
Chuang Mu Xiang	29.11	—
Zi Hua Di Ding	29.91	—
Du Zhong	30.08	—
Zhi Mu	30.10	—
Huang Qin	30.66	—
Wu Mei	31.37	—
Fo Shou	31.71	—
Huang Lian	31.78	—
Yi Mu Cao	33.14	—
Gan Cao	34.61	—
Sang Zhi	35.38	—
Tian Men Dong	35.72	—
Long Yan Rou	36.07	—
Wang Bu Liu Xing	36.41	—
Ze Xie	37.57	—
Chi Shao (Yao)	38.45	—
Mang Xiao	38.71	—
Huang Bai	40.73	—
Fu Shen	40.78	—
Jie Geng	42.01	—
Xi Xian Cao	43.39	—
Ku Shen	46.18	—
Bai Mao Gen	46.36	—
Chuan Lian Zi	46.36	—
Zi Wan	48.14	—
Niu Bang Zi	48.85	—
San Leng	49.11	—
Zi Su Zi	49.98	—
(Shi) Chang Pu	50.00	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/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
62524	62524	0.00	10.06
62523	62523	0.00	10.46
62447	62447	0.00	3.83
62448	62448	0.00	3.92
62332	62332	0.00	12.53
62331	62331	0.00	12.40

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 2.1-2020-01

Validated substance/substance group **Shan Yu Rou**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60226-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G080HS027RH1	62269	40	from supplier
PhytoComm	Shan Yu Rou	G080HS027RH1	62270	40	from supplier
PhytoComm	Shan Yu Rou	G080HS027SK1	62571	40	from supplier
PhytoComm	Shan Yu Rou	G080HS027SK1	62572	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G080HS027RH1	62269 [†]	20
PhytoComm	Shan Yu Rou	G080HS027RH1	62270 [†]	20
PhytoComm	Shan Yu Rou	G080HS027SK1	62571 [†]	20
PhytoComm	Shan Yu Rou	G080HS027SK1	62572 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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	25 641
Type B	0	80	0	13 087
Type C	1	0	4	848

The substance/substance group *Shan Yu Rou* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
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

[†]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 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
Dang Shen	9.29	—
Cang Zhu	9.58	—
Chuan Xiong	10.00	—
San Qi	10.44	—
Lian Qiao	12.53	—
Zhe Bei Mu	13.28	—
Pi Pa Ye	13.36	—
Dang Gui	13.92	—
Zhi Mu	14.05	—
Dang Gui Wei	14.36	—
Yuan Zhi	14.50	—
Mu Dan Pi	14.65	—
Jin Yin Hua	14.69	—
Bai Shao Yao	15.18	—
Gua Lou	15.21	—
Jiao Gu Lan	17.03	—
Gou Qi Zi	17.07	—
Wu Wei Zi	17.16	—
Di Gu Pi	17.93	—
Bai Zi Ren	18.04	—
Chen Pi	18.16	—
Shan Yao	19.10	—
Ren Dong Teng	19.28	—
He Huan Pi	19.39	—
Yu Zhu	19.46	—
Chuang Mu Xiang	19.52	—
Bai Bian Dou	19.75	—
Ye Jiao Teng	20.17	—
Tai Zi Shen	20.24	—
Ji Li	20.42	—
Zhi Ke	22.37	—
Gan Cao	23.28	—
Hou Po	24.28	—
Tu Fu Ling	24.31	—
Lian Zi	24.34	—
Tian Hua Fen	24.47	—
Tao Ren	24.84	—
Huo Ma Ren	25.02	—
Suan Zao Ren	25.28	—
Fu Zi	25.45	—
Fang Ji	26.05	—
Ban Xia (Jiang)	26.75	—
Huang Qin	26.95	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Shen	27.15	—
Ci Wu Jia	27.40	—
Zhu Ru	27.52	—
Bai Zhu	28.22	—
Fo Shou	28.31	—
Mao Dong Qing	29.43	—
Fu Ling	29.51	—
Shen Qu	29.96	—
Da Zao	30.14	—
Rou Gui	30.21	—
Qing Pi	30.36	—
Fu Xiao Mai	30.89	—
Yin Yang Huo	30.95	—
Ji Xue Teng	31.10	—
Jie Geng	31.46	—
Gou Teng	31.48	—
Ma Huang	31.49	—
Long Yan Rou	31.76	—
Yan Hu Suo	31.77	—
Sheng Jiang	31.92	—
Yi Yi Ren	33.90	—
Gui Zhi	33.92	—
Sang Zhi	34.22	—
Mai Men Dong	34.28	—
Hong Jing Tian	34.41	—
Bai Xian Pi	34.68	—
Chai Hu	34.74	—
Che Qian Zi	34.75	—
Tian Men Dong	35.61	—
Chi Shao (Yao)	36.26	—
San Leng	37.04	—
(Shi) Chang Pu	37.40	—
Ling Zhi	37.55	—
Zi Hua Di Ding	38.94	—
Hong Hua	39.21	—
Ze Xie	39.63	—
Cang Er Zi	40.34	—
Qiang Huo	41.06	—
Huang Lian	41.13	—
Mang Xiao	41.30	—
Dan Dou Chi	43.10	—
Du Zhong	44.27	—
Sang Ye	45.84	—
Yu Jin	46.01	—
Zi Wan	46.83	—
Huang Qi	47.92	—
Yi Mu Cao	47.93	—
Wang Bu Liu Xing	48.77	—
Bai Mao Gen	49.82	—

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	11.96
62269	62269	0.00	9.58
62270	62270	0.00	9.29
62572	62572	0.00	12.33

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 2.1-2020-01

Validated substance/substance group **She Gan**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60335-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G048H1034621	61771	40	from supplier
PhytoComm	She Gan	G048H1034621	61828	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G048H1034621	61771 [†]	20
PhytoComm	She Gan	G048H1034621	61828 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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

- 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	10	0	1	842

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.0289 % (> 98.4461 %)	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
Gou Qi Zi	7.20	—
Bai Mao Gen	10.94	—
Dang Gui Wei	11.31	—
Du Huo	12.72	—
Huang Qi	13.61	—
Zi Wan	16.53	—
Chai Hu	16.98	—
(Huai) Niu Xi	17.16	—
Mu Dan Pi	17.93	—
Dang Gui	17.94	—
Chuan Xiong	18.12	—
Cang Zhu	18.52	—
Jie Geng	18.71	—
Sang Zhi	19.31	—
Zhi Mu	19.55	—
San Leng	19.86	—
Bai Shao Yao	19.93	—
Lu Gen	21.02	—
Gui Zhi	22.65	—
Mu Gua	23.55	—
Lian Qiao	23.75	—
Bai Zhu	24.36	—
Long Dan (Cao)	24.54	—
Dang Shen	24.60	—
Chi Shao (Yao)	25.06	—
Fang Feng	26.18	—
Gua Lou	26.60	—
Qing Pi	26.90	—
Yuan Zhi	27.18	—
Di Gu Pi	28.45	—
Ju Hua	29.17	—
Hong Jing Tian	29.30	—
(Shi) Chang Pu	29.51	—
Qin Jiao	30.27	—
Xiang Fu	30.35	—
Mang Xiao	31.43	—
Gan Jiang	31.46	—
Chen Pi	32.69	—
Wu Yao	33.36	—
Ba Ji Tian	33.49	—
Ji Li	34.21	—
Zhi Ke	34.71	—
Ze Xie	35.07	—
Hou Po	35.35	—
Lai Fu Zi	37.78	—
Xie Bai	39.26	—
E Zhu	39.30	—
Yi Mu Cao	39.63	—
Tu Fu Ling	40.31	—
Tian Hua Fen	40.88	—
Ku Shen	42.18	—
Sha Shen (Bei)	42.45	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wei Ling Xian	42.60	—
Xian He Cao	42.82	—
Huo Ma Ren	42.92	—
Bai He	43.06	—
Mai Men Dong	43.27	—
Jiang Huang	44.14	—
Sheng Jiang	44.32	—
Da Zao	44.86	—
Chuan Lian Zi	45.37	—
Niu Bang Zi	45.39	—
Shan Yao	45.89	—
Gan Cao	47.49	—
Ban Zhi Lian	47.78	—
Zi Su Zi	48.01	—
Wu Wei Zi	48.26	—
Huang Lian	49.25	—
He Huan Pi	49.77	—
Jin Yin Hua	49.86	—
Lian Zi	50.13	—

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
61828	61828	0.00	7.34
61771	61771	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 2.1-2020-01

Validated substance/substance group **Shen Qu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60020-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Shen Qu	3	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	G156H0935621	62259	40	from supplier
PhytoComm	Shen Qu	G156H0935621	62260	40	from supplier
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 39 821 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 *Shen Qu*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G156H0935621	62259 [†]	20
PhytoComm	Shen Qu	G156H0935621	62260 [†]	20
PhytoComm	Shen Qu	G156HS222RP1	62459 [†]	20
PhytoComm	Shen Qu	G156HS222RP1	62460 [†]	20
PhytoComm	Shen Qu	G156H0935821	62467 [†]	20
PhytoComm	Shen Qu	G156H0935821	62468 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	g156h0935421	1
PhytoComm	Shen Qu	G156H0935521	1
PhytoComm	Shen Qu	G156H0935121	1
PhytoComm	Shen Qu	G156H0935521	1
PhytoComm	Shen Qu	G156H0935522	1

- 848 spectra from 12 *Apo-Ident* customers from a total of 512 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	0	240	0	25 561
Type B	0	120	0	13 047
Type C	1	1	4	847

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	100.0000 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	99.9070 % (> 99.3190 %)	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 *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	4.81	—
Bai Bian Dou	5.90	—
Ren Dong Teng	7.34	—
He Huan Pi	7.46	—
Lian Zi	8.63	—
Tian Hua Fen	8.96	—
Shan Yao	9.04	—
Ji Li	9.47	—
Tai Zi Shen	9.52	—
Pi Pa Ye	9.71	—
Tao Ren	9.92	—
Bai Shao Yao	10.06	—
Zhe Bei Mu	10.11	—
Ling Zhi	10.31	—
Lian Qiao	10.66	—
Ban Xia (Jiang)	10.72	—
Chuan Xiong	11.24	—
Fu Zi	11.32	—
Bai Zi Ren	11.86	—
Gua Lou	12.25	—
Gou Teng	12.62	—
Ji Xue Teng	12.70	—
Ye Jiao Teng	12.97	—
Sheng Jiang	13.32	—
Ci Wu Jia	13.42	—
Chen Pi	13.63	—
Rou Gui	13.84	—
Bai Xian Pi	13.86	—
Gui Zhi	14.02	—
Fu Ling	14.35	—
Di Gu Pi	14.56	—
Fu Xiao Mai	14.65	—
Dan Shen	15.09	—
Cang Er Zi	15.98	—
Yan Hu Suo	16.66	—
Yuan Zhi	17.17	—
Yu Jin	17.91	—
Yin Yang Huo	18.29	—
Mao Dong Qing	18.32	—
Dan Dou Chi	18.45	—
Hou Po	18.78	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tu Fu Ling	18.92	—
Jiao Gu Lan	18.92	—
Yi Yi Ren	19.06	—
Zhu Ru	19.13	—
Ma Huang Gen	19.24	—
Jin Yin Hua	19.78	—
Ma Huang	20.25	—
Che Qian Zi	20.53	—
Dong Kui Zi	20.59	—
Chai Hu	20.72	—
Hong Jing Tian	20.77	—
Fo Shou	21.15	—
San Qi	21.55	—
Cang Zhu	22.11	—
Yu Zhu	22.31	—
Da Zao	22.80	—
Fang Ji	22.90	—
Dang Gui Wei	23.29	—
Wu Wei Zi	25.64	—
Tian Men Dong	25.92	—
Zhi Ke	27.53	—
Dang Shen	27.91	—
Dang Gui	28.21	—
Shan Yu Rou	28.23	—
Qiang Huo	28.34	—
Gou Qi Zi	29.00	—
Long Yan Rou	29.48	—
Zhi Mu	29.54	—
Qing Pi	30.33	—
Wang Bu Liu Xing	30.72	—
Fu Shen	31.50	—
Mu Dan Pi	31.51	—
Zhi Zi	31.77	—
Suan Zao Ren	31.84	—
Huang Qin	32.42	—
Huang Lian	33.48	—
Du Zhong	35.37	—
Sang Zhi	36.26	—
Mang Xiao	37.58	—
Ze Xie	38.33	—
Bai Zhu	39.01	—
Zi Hua Di Ding	39.31	—
Jie Geng	39.45	—
Chuang Mu Xiang	39.51	—
Gan Cao	39.60	—
Chi Shao (Yao)	40.48	—
Jiang Huang	45.50	—
Ku Shen	45.78	—
Chuan Lian Zi	46.24	—
Nü Zhen Zi	48.83	—
Yi Mu Cao	50.24	—

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
62460	62460	0.00	4.81
62459	62459	0.00	5.19
62467	62467	0.00	16.02
62468	62468	0.00	15.98
62259	62259	0.00	25.84
62260	62260	0.00	25.33

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 2.1-2020-01

Validated substance/substance group **Sheng Jiang**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60017-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G253H0534623	62141	40	from supplier
PhytoComm	Sheng Jiang	G253H0534623	62142	40	from supplier
PhytoComm	Sheng Jiang	G253HS085SK1	62549	40	from supplier
PhytoComm	Sheng Jiang	G253HS085SK1	62550	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G253H0534623	62141 [†]	20
PhytoComm	Sheng Jiang	G253H0534623	62142 [†]	20
PhytoComm	Sheng Jiang	G253HS085SK1	62549 [†]	20
PhytoComm	Sheng Jiang	G253HS085SK1	62550 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G253H0534522	1
PhytoComm	Sheng Jiang	G253H0534321	1
PhytoComm	Sheng Jiang	G253H0534522	3

- 847 spectra from 12 *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 *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	152	8	25 635
Type B	19	70	10	13 068
Type C	1	0	6	846

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	99.9782 % (> 99.9633 %)	95.0000 % (> 93.1250 %)
Type B	99.8263 % (> 99.7974 %)	87.5000 % (> 83.7500 %)
Type C	99.8837 % (> 99.2955 %)	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
Zhu Ru	2.47	—
Ban Xia (Jiang)	3.62	—
Di Gu Pi	4.20	—
Gui Zhi	4.95	—
Rou Gui	5.37	—
Fu Xiao Mai	5.72	—
Yi Yi Ren	6.62	—
Ci Wu Jia	8.08	—
Tian Hua Fen	8.67	—
Fu Ling	9.03	—
He Huan Pi	9.06	—
Bai Bian Dou	10.34	—
Tao Ren	11.04	—
Zhe Bei Mu	11.75	—
Sang Zhi	12.00	—
Lian Zi	12.85	—
Bai Shao Yao	12.87	—
Gou Teng	13.43	—
Shen Qu	13.47	—
Ji Li	13.61	—
Bai Zi Ren	13.76	—
San Qi	13.78	—
Wang Bu Liu Xing	15.28	—
Ling Zhi	15.85	—
Tai Zi Shen	16.24	—
Huo Ma Ren	16.44	—
Ji Xue Teng	16.51	—
Bai Xian Pi	17.99	—
Ma Huang Gen	18.83	—
Jie Geng	18.89	—
Bai He	18.96	—
Chuan Lian Zi	19.17	—
E Zhu	19.26	—
Mai Ya	20.18	—
Lai Fu Zi	20.99	—
Fu Shen	21.28	—
Gua Lou	21.51	—
Gan Jiang	22.00	—
Fo Shou	22.15	—
Ye Jiao Teng	22.76	—
Bai Zhu	23.59	—
Yuan Zhi	23.65	—
Long Yan Rou	23.74	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Long Dan (Cao)	23.94	—
Xiang Fu	23.98	—
Tian Men Dong	24.08	—
Da Zao	24.10	—
Pi Pa Ye	24.30	—
Shan Yao	24.77	—
Qing Pi	24.94	—
Lian Qiao	25.02	—
Mai Men Dong	25.11	—
Dang Shen	25.11	—
Yu Jin	25.47	—
Yan Hu Suo	25.54	—
San Leng	26.01	—
Zi Su Zi	26.08	—
Chen Pi	26.43	—
Mang Xiao	26.61	—
(Shi) Chang Pu	26.67	—
Mu Dan Pi	27.49	—
She Gan	27.56	—
Chuan Xiong	28.14	—
Ren Dong Teng	28.22	—
Qin Jiao	28.49	—
Ban Zhi Lian	28.94	—
Dang Gui Wei	29.14	—
Jiang Huang	29.19	—
Du Huo	29.35	—
Dang Gui	29.69	—
Tu Fu Ling	29.84	—
Gou Qi Zi	30.02	—
Lu Gen	30.79	—
Dan Shen	32.09	—
Cang Er Zi	32.37	—
Huang Qi	32.78	—
Yi Mu Cao	33.17	—
Zhi Mu	33.17	—
Fu Zi	33.23	—
Shan Yu Rou	33.67	—
Dan Dou Chi	33.73	—
Cang Zhu	33.90	—
Fang Ji	33.99	—
Wu Wei Zi	34.20	—
Ba Ji Tian	34.33	—
Chai Hu	34.97	—
Jiao Gu Lan	35.39	—
Sha Shen (Bei)	35.42	—
Bai Mao Gen	35.45	—
Mao Dong Qing	35.56	—
Qiang Huo	35.66	—
Mu Gua	36.28	—
Zhi Ke	36.57	—
Huang Lian	36.88	—
Jin Yin Hua	37.04	—
Yin Yang Huo	37.08	—
(Bai) Dou Kou	37.13	—
Mo Yao	37.14	—
Zhu Ling	37.37	—
Zi Wan	37.48	—
Huang Qin	37.62	—
(Huai) Niu Xi	37.68	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ze Xie	38.33	—
Xie Bai	38.41	—
Chi Shao (Yao)	39.65	—
Yu Zhu	40.67	—
Ku Shen	40.77	—
Du Zhong	41.15	—
Ma Huang	42.07	—
Hou Po	42.10	—
Fang Feng	42.24	—
Pu Gong Ying	42.56	—
Niu Bang Zi	42.65	—
Che Qian Zi	43.53	—
Sha Ren	43.75	—
Wu Yao	44.58	—
Hong Jing Tian	47.71	—

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
62142	62142	0.00	8.67
62549	62549	0.00	2.47
62141	62141	0.00	10.17
62550	62550	0.00	2.95

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 2.1-2020-01

Validated substance/substance group **Shi Gao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50368-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Shi Gao	1	1	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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.

- 130 spectra of 4 reference samples 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. 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	G120H0541721	62229	45
PhytoComm	Shi Gao	G120H0541721	62230	45
PhytoComm	Shi Gao	G120HS081SH1	62561 [†]	20
PhytoComm	Shi Gao	G120HS081SH1	62562 [†]	20

- 13 037 spectra from a total of 318 batches from further 172 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

- 850 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 *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	25 721
Type B	0	40	90	13 037
Type C	0	0	3	850

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9421 %)	30.7692 % (> 28.4615 %)
Type C	100.0000 % (> 98.8257 %)	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	53.62	—

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	54.10
62562	62562	0.00	53.62

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 2.1-2020-01

Validated substance/substance group **Shu Di (Huang)**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60006-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Shu Di (Huang)	2	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)	G210H1601822	62297	40	from supplier
PhytoComm	Shu Di (Huang)	G210H1601822	62298	40	from supplier
PhytoComm	Shu Di (Huang)	G210H1601823	62493	40	from supplier
PhytoComm	Shu Di (Huang)	G210H1601823	62494	40	from supplier

Validation samples

A total of 39 821 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 *Shu Di (Huang)*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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)	G210H1601822	62297 [†]	20
PhytoComm	Shu Di (Huang)	G210H1601822	62298 [†]	20
PhytoComm	Shu Di (Huang)	G210H1601823	62493 [†]	20
PhytoComm	Shu Di (Huang)	G210H1601823	62494 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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)	G210H1601221	1
PhytoComm	Shu Di (Huang)	G210H1601421	6
PhytoComm	Shu Di (Huang)	G210H1601521	3
PhytoComm	Shu Di (Huang)	G210H1601121	1

- 838 spectra from 12 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	80	0	13 087
Type C	6	0	15	832

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.4574 % (> 98.8685 %)	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 *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
Ai Ye	9.11	—
Jue Ming Zi	15.61	—
Yu Xing Cao	20.31	—
(Sheng) Di Huang	23.00	—
Guang Huo Xiang	27.83	—
Pu Gong Ying	29.51	—
Jing Jie	41.76	—
Hua Shi	42.87	—
Wu Jia Pi	52.10	—

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
62494	62494	0.00	9.11
62493	62493	0.00	9.80
62297	62297	0.00	14.31
62298	62298	0.00	13.83

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 2.1-2020-01

Validated substance/substance group **Suan Zao Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60059-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G254H1418621	61893	40	from supplier
PhytoComm	Suan Zao Ren	G254H1418621	61894	40	from supplier
PhytoComm	Suan Zao Ren	G254H1418622	62133	40	from supplier
PhytoComm	Suan Zao Ren	G254H1418622	62134	40	from supplier
PhytoComm	Suan Zao Ren	G254HS310SM1	62717	40	from supplier
PhytoComm	Suan Zao Ren	G254HS310SM1	62718	40	from supplier

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G254H1418621	61893 [†]	20
PhytoComm	Suan Zao Ren	G254H1418621	61894 [†]	20
PhytoComm	Suan Zao Ren	G254H1418622	62133 [†]	20
PhytoComm	Suan Zao Ren	G254H1418622	62134 [†]	20
PhytoComm	Suan Zao Ren	G254HS310SM1	62717 [†]	20
PhytoComm	Suan Zao Ren	G254HS310SM1	62718 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 850 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 *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	25 561
Type B	2	120	0	13 045
Type C	1	0	3	849

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	99.9903 % (> 99.9614 %)	100.0000 % (> 95.0000 %)
Type C	99.8837 % (> 99.2966 %)	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
Yuan Zhi	7.42	—
Lian Qiao	7.79	—
Bai Bian Dou	9.38	—
Cang Zhu	11.42	—
Ji Li	11.60	—
Bai Zi Ren	12.11	—
Jiao Gu Lan	12.86	—
Shan Yao	13.46	—
Mao Dong Qing	13.50	—
Lian Zi	14.32	—
Ban Xia (Jiang)	14.32	—
San Qi	15.15	—
Ling Zhi	15.89	—
Chen Pi	18.21	—
Dang Shen	18.55	—
Bai Shao Yao	19.29	—
Shen Qu	19.75	—
Dong Gua Zi	20.06	—
Pi Pa Ye	20.08	—
Jin Yin Hua	20.14	—
Huo Ma Ren	20.20	—
Dan Shen	20.35	—
Wu Wei Zi	20.62	—
Zhe Bei Mu	20.63	—
Shan Yu Rou	20.65	—
Chuan Xiong	21.29	—
Fu Xiao Mai	21.53	—
Gou Qi Zi	21.84	—
Yu Zhu	22.17	—
Chuang Mu Xiang	22.64	—
Mu Dan Pi	22.93	—
Rou Gui	23.09	—
Hou Po	23.61	—
Gua Lou	24.00	—
Tai Zi Shen	24.07	—
Fu Zi	24.28	—
Fu Ling	24.70	—
Sheng Jiang	24.72	—
Di Gu Pi	25.01	—
Chai Hu	25.24	—
Tu Fu Ling	25.74	—
Gui Zhi	26.69	—
Fang Ji	26.93	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
He Huan Pi	27.08	—
Bai Xian Pi	27.10	—
Ren Dong Teng	27.18	—
Zhu Ru	27.20	—
Cang Er Zi	27.63	—
Dang Gui	27.83	—
Dang Gui Wei	28.21	—
Zhi Mu	29.80	—
Mang Xiao	29.94	—
Dan Dou Chi	29.94	—
Zhi Ke	30.30	—
Yan Hu Suo	30.38	—
Tian Hua Fen	30.41	—
Ye Jiao Teng	31.00	—
Xie Bai	31.02	—
Qing Pi	31.11	—
Hong Jing Tian	32.33	—
Bai Zhu	32.35	—
Du Zhong	32.54	—
Zi Su Zi	32.99	—
Yin Yang Huo	33.17	—
Gou Teng	33.27	—
Yu Jin	33.40	—
Yi Yi Ren	33.66	—
Ji Xue Teng	34.08	—
Ci Wu Jia	34.21	—
Tao Ren	34.40	—
She Gan	34.94	—
Wang Bu Liu Xing	34.99	—
Zhi Zi	36.01	—
Qiang Huo	36.51	—
Che Qian Zi	37.89	—
Huang Qin	38.29	—
Zi Hua Di Ding	38.35	—
Da Zao	39.80	—
Tian Men Dong	40.64	—
Chi Shao (Yao)	41.40	—
Chuan Lian Zi	41.42	—
Sang Zhi	41.77	—
Wei Ling Xian	42.04	—
Ze Xie	42.12	—
Xu Duan	42.64	—
Ma Huang	43.15	—
Long Yan Rou	43.74	—
Jiang Huang	43.78	—
Gan Cao	44.22	—
Huang Lian	44.29	—
Ma Huang Gen	44.86	—
Gan Jiang	47.14	—
Sang Ye	47.92	—
Ku Shen	48.25	—
Jie Geng	49.19	—

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
61893	61893	0.00	21.06
62718	62718	0.00	7.83
61894	61894	0.00	20.06
62717	62717	0.00	7.42
62133	62133	0.00	21.29
62134	62134	0.00	22.35

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 2.1-2020-01

Validated substance/substance group **Tai Zi Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60114-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	1	80	0	25 720
Type B	0	40	0	13 127
Type C	0	0	2	851

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	99.9979 % (> 99.9831 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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	5.18	—
Zhe Bei Mu	6.07	—
Shen Qu	7.00	—
He Huan Pi	7.34	—
Ban Xia (Jiang)	8.55	—
Tian Hua Fen	9.05	—
Ye Jiao Teng	9.36	—
Tao Ren	9.52	—
Lian Zi	9.52	—
Huo Ma Ren	9.71	—
Ji Li	9.86	—
Ji Xue Teng	9.88	—
Ci Wu Jia	10.66	—
Gui Zhi	10.71	—
Gou Teng	11.14	—
Di Gu Pi	11.25	—
Rou Gui	11.33	—
Bai Bian Dou	11.57	—
Fo Shou	11.62	—
Gua Lou	11.68	—
Bai Zi Ren	12.32	—
Pi Pa Ye	12.35	—
Sheng Jiang	12.94	—
Lian Qiao	13.23	—
Chen Pi	15.33	—
Zhu Ru	15.97	—
Da Zao	16.50	—
Ling Zhi	16.73	—
Shan Yao	16.98	—
Yan Hu Suo	17.05	—
Fu Zi	17.35	—
Fu Ling	17.39	—
Ren Dong Teng	17.78	—
Yi Yi Ren	18.13	—
Fu Xiao Mai	18.20	—
Yuan Zhi	18.20	—
Bai Xian Pi	18.25	—
Ma Huang Gen	18.53	—
Chuan Xiong	19.00	—
Long Yan Rou	19.65	—
Dang Gui Wei	20.03	—
Tian Men Dong	20.28	—
Jin Yin Hua	20.40	—
Ma Huang	20.65	—
Yu Zhu	21.72	—
Fu Shen	22.39	—
Dan Shen	22.64	—
Jiao Gu Lan	23.02	—
Dang Shen	23.30	—
Hou Po	23.38	—
Tu Fu Ling	23.64	—
Yin Yang Huo	24.31	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yu Jin	24.73	—
Cang Zhu	25.36	—
Shan Yu Rou	25.41	—
Zhi Ke	25.51	—
Dan Dou Chi	25.74	—
Dang Gui	25.90	—
Zhi Mu	26.00	—
San Qi	26.07	—
Fang Ji	26.14	—
Huang Qin	27.31	—
Wu Wei Zi	27.65	—
Mao Dong Qing	28.31	—
Che Qian Zi	28.77	—
Hong Jing Tian	31.11	—
Gou Qi Zi	31.77	—
Chai Hu	32.51	—
Wang Bu Liu Xing	32.64	—
Gan Cao	36.09	—
Bai Zhu	36.65	—
Qiang Huo	36.70	—
Huang Lian	37.47	—
Sang Zhi	37.49	—
Mang Xiao	38.50	—
Mu Dan Pi	39.42	—
Qing Pi	39.55	—
Jie Geng	39.96	—
Ze Xie	40.83	—
Chi Shao (Yao)	42.24	—
Du Zhong	45.10	—
Suan Zao Ren	45.15	—
Chuang Mu Xiang	46.88	—
San Leng	47.81	—
Zi Hua Di Ding	47.96	—

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
62596	62596	0.00	5.18
62595	62595	0.00	6.05

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 2.1-2020-01

Validated substance/substance group **Tao Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60454-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	0	78	2	25 721
Type B	0	34	6	13 127
Type C	0	0	2	851

The substance/substance group *Tao 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 % (> 99.9703 %)	97.5000 % (> 93.7500 %)
Type B	100.0000 % (> 99.9424 %)	85.0000 % (> 77.5000 %)
Type C	100.0000 % (> 98.8279 %)	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	4.09	—
Gui Zhi	5.64	—
Ban Xia (Jiang)	5.68	—
Lian Zi	6.38	—
Ci Wu Jia	6.75	—
Sheng Jiang	6.90	—
Rou Gui	9.52	—
Bai Bian Dou	9.70	—
Huo Ma Ren	10.18	—
Bai Zi Ren	10.74	—
Di Gu Pi	11.05	—
Bai Shao Yao	11.10	—
Yi Yi Ren	11.29	—
Tai Zi Shen	11.36	—
Gou Teng	11.63	—
Bai Xian Pi	11.97	—
Ji Li	13.23	—
Fo Shou	13.27	—
Shen Qu	13.39	—
Fu Ling	13.60	—
Ye Jiao Teng	13.67	—
Zhu Ru	13.76	—
Ling Zhi	13.95	—
Zhe Bei Mu	15.14	—
Fu Xiao Mai	15.72	—
Gua Lou	16.18	—
Ma Huang Gen	16.59	—
Chen Pi	16.93	—
Tian Men Dong	16.99	—
Ji Xue Teng	17.06	—
Tian Hua Fen	17.13	—
Tu Fu Ling	18.12	—
Ren Dong Teng	19.92	—
Chuan Xiong	20.42	—
Yuan Zhi	21.41	—
Long Yan Rou	21.50	—
Pi Pa Ye	21.59	—
Da Zao	23.23	—
Shan Yao	23.25	—
Fu Shen	23.50	—
Dan Shen	23.89	—
Yu Jin	24.02	—
Lian Qiao	24.04	—
Dan Dou Chi	25.43	—
Mao Dong Qing	25.58	—
Dang Shen	25.59	—
Wang Bu Liu Xing	25.82	—
Yan Hu Suo	27.46	—
Yu Zhu	28.17	—
Fu Zi	28.86	—
Ma Huang	29.05	—
Zhi Mu	29.08	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
San Qi	29.66	—
Dang Gui Wei	29.98	—
Jiao Gu Lan	30.04	—
Yin Yang Huo	30.23	—
Hong Jing Tian	31.09	—
Zhi Ke	31.25	—
Chai Hu	31.78	—
Shan Yu Rou	32.46	—
Cang Zhu	32.93	—
Dang Gui	34.81	—
Fang Ji	35.02	—
Hou Po	36.71	—
Jin Yin Hua	37.33	—
Wu Wei Zi	37.56	—
Mang Xiao	37.70	—
Huang Qin	38.21	—
Che Qian Zi	38.45	—
Gou Qi Zi	42.50	—
Qiang Huo	43.12	—
Suan Zao Ren	43.43	—
Sang Zhi	43.74	—
Bai Zhu	44.05	—
Jie Geng	44.85	—
Huang Lian	45.02	—
Ze Xie	45.09	—
Chuan Lian Zi	47.44	—
Qing Pi	48.00	—

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
62610	62610	0.00	4.29
62609	62609	0.00	4.09

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 2.1-2020-01

Validated substance/substance group **Tian Hua Fen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60139-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Tian Hua Fen	2	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

Validation samples

A total of 39 821 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 *Tian Hua Fen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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	h0417022	1
Phytocomm	Tian Hua Fen	G242H0417522	1
PhytoComm	Tian Hua Fen	G242H0417522	1

- 848 spectra from 12 *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 *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	0	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	5	848

The substance/substance group *Tian Hua Fen* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8240 %)	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 *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
Tai Zi Shen	6.65	—
Shen Qu	8.46	—
Huo Ma Ren	8.56	—
Bai Bian Dou	8.82	—
Sheng Jiang	8.92	—
Zhe Bei Mu	9.10	—
Ji Li	9.18	—
Sang Zhi	9.56	—
Pi Pa Ye	9.61	—
Bai Shao Yao	9.68	—
Bai Zi Ren	9.69	—
Fu Zi	9.77	—
Lian Qiao	9.80	—
He Huan Pi	10.16	—
Shan Yao	11.31	—
Ji Xue Teng	11.39	—
Lian Zi	12.09	—
Ren Dong Teng	12.72	—
Gua Lou	12.81	—
Gui Zhi	12.86	—
Di Gu Pi	13.31	—
Tao Ren	13.49	—
Yan Hu Suo	13.55	—
Rou Gui	13.85	—
Gan Jiang	14.15	—
Jin Yin Hua	14.29	—
Ban Xia (Jiang)	14.94	—
Ye Jiao Teng	14.96	—
Hou Po	15.19	—
Ling Zhi	15.24	—
Gou Teng	15.29	—
Ci Wu Jia	15.56	—
Xiang Fu	15.67	—
Chen Pi	16.72	—
Fu Xiao Mai	16.90	—
Jie Geng	16.97	—
Chuan Xiong	17.27	—
Bai Xian Pi	17.95	—
San Qi	18.57	—
Zhu Ru	18.61	—
Ma Huang	18.67	—
(Shi) Chang Pu	18.70	—
Fu Ling	19.19	—
Jiao Gu Lan	19.38	—
Yi Yi Ren	19.60	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dan Shen	19.84	—
Yuan Zhi	20.02	—
Cang Zhu	20.26	—
Qing Pi	20.36	—
Yu Jin	20.44	—
Che Qian Zi	20.51	—
Da Zao	20.90	—
Lai Fu Zi	20.95	—
Fang Ji	21.09	—
Long Dan (Cao)	21.48	—
Dan Dou Chi	21.55	—
Chuan Lian Zi	21.66	—
Yin Yang Huo	21.99	—
Wu Wei Zi	22.26	—
Yu Zhu	22.35	—
Sha Shen (Bei)	22.37	—
San Leng	22.60	—
Dang Gui Wei	22.62	—
E Zhu	22.74	—
Ma Huang Gen	22.78	—
Dang Shen	22.91	—
Fo Shou	23.07	—
Dang Gui	23.10	—
She Gan	23.31	—
Mao Dong Qing	24.05	—
Du Huo	24.76	—
Bai Zhu	24.96	—
Shan Yu Rou	24.98	—
Bai He	25.12	—
Cang Er Zi	25.14	—
Zhi Ke	25.18	—
Yi Mu Cao	25.60	—
Mai Ya	25.79	—
Gou Qi Zi	25.91	—
Lu Gen	26.47	—
Chai Hu	26.49	—
Jiang Huang	26.67	—
Hong Jing Tian	26.82	—
Zi Su Zi	27.01	—
Mai Men Dong	27.26	—
Tu Fu Ling	27.54	—
Mang Xiao	27.56	—
Huang Qi	27.80	—
Qin Jiao	28.06	—
Huang Qin	28.11	—
(Huai) Niu Xi	28.87	—
Zi Wan	29.14	—
Bai Mao Gen	29.19	—
Long Yan Rou	29.29	—
Mu Dan Pi	29.50	—
Tian Men Dong	30.04	—
Zhi Mu	30.19	—
Qiang Huo	30.71	—
(Bai) Dou Kou	31.17	—
Fu Shen	31.99	—
Ba Ji Tian	32.01	—
Ban Zhi Lian	32.32	—
Ze Xie	32.32	—
Wang Bu Liu Xing	32.33	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Lian	34.02	—
Ku Shen	34.03	—
Gan Cao	34.78	—
Mu Gua	34.94	—
Niu Bang Zi	35.53	—
Suan Zao Ren	36.15	—
Pu Gong Ying	36.80	—
Fang Feng	36.96	—
Zhu Ling	37.10	—
Chi Shao (Yao)	37.18	—
Wu Yao	37.31	—
Du Zhong	37.41	—
Xie Bai	37.45	—
Sha Ren	38.76	—
Chuang Mu Xiang	39.12	—
Zi Hua Di Ding	39.32	—
Mo Yao	44.00	—
Xian He Cao	46.79	—
Wei Ling Xian	48.91	—
Dong Gua Zi	49.60	—

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
62634	62634	0.00	6.65
62316	62316	0.00	8.94
62633	62633	0.00	6.74
62315	62315	0.00	8.92

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 2.1-2020-01

Validated substance/substance group **Tian Men Dong**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60156-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Tian Men Dong; Asparagi cochinensis radix

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Tian Men Dong	1	0	3

Second-stage model

For differentiation of the substance/substance group *Tian 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 *Tian Men Dong*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Tian Men Dong	G037HS041RR1	62437	40	from supplier
PhytoComm	Tian Men Dong	G037HS041RR1	62438	40	from supplier

Validation samples

A total of 39 821 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 *Tian Men Dong*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Tian 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	Tian Men Dong	G037HS041RR1	62437 [†]	20
PhytoComm	Tian Men Dong	G037HS041RR1	62438 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Tian Men Dong*.
- 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 Men Dong	G037H0415221	1
Phytocomm	Tian Men Dong	g03h0415521	1
Phytocomm	Tian Men Dong	g037h0415221	1
Phytocomm	Tian Men Dong	G037H0415321	1

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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 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 *Tian 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	25 721
Type B	0	39	1	13 127
Type C	0	0	4	849

The substance/substance group *Tian 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8246 %)	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 *Tian 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
Fo Shou	8.20	—
Dang Shen	8.42	—
Chen Pi	8.77	—
Yuan Zhi	11.44	—
Zhi Mu	11.75	—
Da Zao	14.09	—
Gua Lou	14.69	—
Ye Jiao Teng	15.12	—
Rou Gui	15.36	—
Chuan Xiong	15.47	—
Di Gu Pi	15.88	—
Bai Shao Yao	15.92	—
Mai Men Dong	16.93	—
Pi Pa Ye	18.05	—
Shen Qu	19.70	—
He Huan Pi	19.76	—
Dang Gui Wei	20.11	—
Fu Ling	20.54	—
San Qi	21.72	—
Lian Zi	21.84	—
Ji Li	22.11	—
Fu Xiao Mai	22.30	—
Gui Zhi	22.73	—
Dang Gui	22.86	—
Shan Yao	23.25	—
Bai Bian Dou	24.30	—
Cang Zhu	24.32	—
Tai Zi Shen	24.45	—
Jie Geng	24.64	—
Ban Xia (Jiang)	24.66	—
Shan Yu Rou	24.68	—
Tu Fu Ling	24.84	—
Yu Zhu	25.08	—
Bai Zi Ren	25.32	—
Ci Wu Jia	25.54	—
Tao Ren	25.61	—
Zhu Ru	26.30	—
Bai Xian Pi	27.09	—
Sheng Jiang	27.36	—
Ren Dong Teng	28.42	—
Gou Qi Zi	28.55	—
Long Yan Rou	28.87	—
Bai Zhu	29.36	—
Lian Qiao	30.64	—
Gou Teng	30.76	—
Chuang Mu Xiang	31.56	—
Zhe Bei Mu	32.20	—
Tian Hua Fen	32.79	—
Ji Xue Teng	33.05	—
Ling Zhi	33.51	—
Huo Ma Ren	34.22	—
Zhi Ke	34.32	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jiao Gu Lan	34.85	—
Huang Qin	34.86	—
Ma Huang	35.55	—
Mao Dong Qing	36.29	—
Dan Shen	36.83	—
Yi Yi Ren	37.01	—
Wang Bu Liu Xing	37.01	—
Ze Xie	37.36	—
Fu Shen	37.80	—
Suan Zao Ren	37.98	—
Gan Cao	38.52	—
Wu Wei Zi	40.19	—
Mang Xiao	40.65	—
Fu Zi	41.76	—
Jin Yin Hua	42.04	—
Yin Yang Huo	42.28	—
Ma Huang Gen	42.75	—
Huang Qi	42.91	—
Hou Po	44.26	—
Hong Jing Tian	44.69	—
Chuan Lian Zi	45.03	—
Sang Zhi	45.34	—
Fang Ji	45.34	—
Yan Hu Suo	45.71	—
San Leng	47.30	—
Che Qian Zi	47.85	—
Dan Dou Chi	49.38	—
Yu Jin	49.50	—
Chi Shao (Yao)	49.76	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Tian 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
62437	62437	0.00	8.42
62438	62438	0.00	8.20

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 2.1-2020-01

Validated substance/substance group **Tu Fu Ling**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50885-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G231H0325621	61915	40	from supplier
PhytoComm	Tu Fu Ling	G231H0325621	61916	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G231H0325621	61915 [†]	20
PhytoComm	Tu Fu Ling	G231H0325621	61916 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	G231H0325421	2
PhytoComm	Tu Fu Ling	G231H0325121	1
PhytoComm	Tu Fu Ling	G23H0325121	1
PhytoComm	Tu Fu Ling	G231H0325421	1

- 847 spectra from 12 *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 *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	25 561
Type B	0	117	3	13 047
Type C	1	1	5	846

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	97.5000 % (> 95.0000 %)
Type C	99.8837 % (> 99.2955 %)	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
Bai Xian Pi	5.48	—
Xi Xian Cao	8.37	—
Ye Jiao Teng	8.80	—
Dan Dou Chi	9.37	—
Rou Gui	12.60	—
Ji Xue Teng	12.88	—
Du Zhong	13.68	—
E Zhu	14.13	—
Lian Zi	14.38	—
Mao Dong Qing	14.60	—
Ling Zhi	14.74	—
Pu Gong Ying	14.87	—
Yin Yang Huo	15.23	—
Yuan Zhi	15.45	—
Bai Bian Dou	15.90	—
Chen Pi	16.73	—
Ren Dong Teng	16.90	—
Zhi Ke	17.47	—
Xin Yi	17.97	—
Gui Zhi	18.07	—
Xia Ku Cao	18.31	—
Ji Li	18.47	—
Jiao Gu Lan	18.72	—
Yu Jin	19.04	—
Shan Yao	19.12	—
Fang Ji	19.35	—
Gou Teng	19.62	—
Sha Shen (Bei)	19.65	—
Sang Zhi	19.88	—
Jiang Huang	19.92	—
Dan Shen	20.28	—
Huang Lian	20.38	—
Mang Xiao	20.70	—
Chuan Xiong	20.77	—
Ma Huang Gen	20.82	—
Nü Zhen Zi	21.14	—
He Huan Pi	21.33	—
Qiang Huo	21.39	—
Hou Po	21.46	—
Shen Qu	21.65	—
Sang Ye	21.84	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Jin Yin Hua	22.05	—
Bai Shao Yao	22.97	—
Huo Ma Ren	23.28	—
Chai Hu	23.56	—
Fo Shou	23.65	—
Mu Dan Pi	24.50	—
Ma Huang	24.71	—
Di Gu Pi	24.95	—
Sheng Jiang	25.05	—
Xian He Cao	25.09	—
Sha Ren	25.11	—
Fu Ling	25.39	—
Long Dan (Cao)	25.39	—
Zhu Ling	25.48	—
Ban Zhi Lian	25.74	—
Ban Xia (Jiang)	26.02	—
(Bai) Dou Kou	26.18	—
Yan Hu Suo	26.55	—
Che Qian Zi	27.09	—
Pi Pa Ye	27.76	—
Long Yan Rou	28.05	—
Lian Qiao	28.43	—
San Leng	28.56	—
Fu Zi	28.72	—
Gua Lou	28.94	—
Zi Hua Di Ding	29.03	—
Yi Yi Ren	29.12	—
Tian Men Dong	29.50	—
Cang Er Zi	29.63	—
Tao Ren	29.81	—
Hu Zhang	29.91	—
Fu Xiao Mai	30.84	—
Mai Ya	30.88	—
Da Zao	31.05	—
Qing Pi	31.36	—
Fu Shen	31.63	—
Du Huo	32.05	—
Wu Mei	32.13	—
Chuan Lian Zi	32.17	—
Ba Ji Tian	32.22	—
Tai Zi Shen	32.69	—
Gou Qi Zi	33.46	—
Zhe Bei Mu	33.57	—
Bai He	33.64	—
Ci Wu Jia	34.32	—
Hong Jing Tian	34.34	—
San Qi	34.39	—
Cang Zhu	34.57	—
Wu Yao	34.72	—
Jie Geng	34.83	—
She Gan	34.87	—
Wu Wei Zi	35.52	—
Gan Jiang	35.57	—
Fang Feng	36.42	—
Yi Mu Cao	36.60	—
Chi Shao (Yao)	36.81	—
Lu Gen	36.88	—
Tian Hua Fen	37.21	—
Wang Bu Liu Xing	37.51	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dang Gui	37.81	—
Bai Hua She She Cao	38.71	—
Xiang Fu	38.96	—
(Huai) Niu Xi	39.09	—
Dang Gui Wei	39.17	—
Bai Zhu	39.90	—
Zhu Ru	40.00	—
Zhi Mu	40.10	—
Zi Su Zi	40.10	—
Han Lian Cao	40.42	—
Mo Yao	41.33	—
Huang Bai	41.49	—
Ze Xie	42.08	—
Sang Ji Shend	42.44	—
Niu Bang Zi	42.72	—
Lai Fu Zi	42.83	—
Yu Zhu	43.32	—
Bu Gu Zhi	44.14	—
Qin Jiao	44.20	—
Huang Qin	44.62	—
Huang Jing	44.85	—
Bai Mao Gen	45.00	—
He Shou Wu	45.51	—
Bai Zi Ren	46.08	—
Dang Shen	46.10	—
(Shi) Chang Pu	46.13	—
Huang Qi	46.46	—
Shan Yu Rou	46.62	—
Xie Bai	46.63	—
Wei Ling Xian	46.98	—
Zi Wan	48.00	—
Sang Bai Pi	48.30	—
Hong Hua	50.09	—

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
62641	62641	0.00	8.37
62427	62427	0.00	5.48
61915	61915	0.00	14.40
62428	62428	0.00	6.83
61916	61916	0.00	14.13
62642	62642	0.00	9.99

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 2.1-2020-01

Validated substance/substance group **Tu Si Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60217-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Tu Si Zi	1	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

Validation samples

A total of 39 821 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 *Tu Si Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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
Fagron	Tu Si Zi	G086H1243521	2
Phytocomm	Tu Si Zi	h1243022	1
Phytocomm	Tu Si Zi	G086H1243521	1
PhytoComm	Tu Si Zi	G086H1243521	1

- 848 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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 *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
Yu Xing Cao	24.78	—
Bu Gu Zhi	26.06	—
Mang Xiao	30.85	—
Dan Zhu Ye	32.01	—
Sha Ren	32.64	—
(Sheng) Di Huang	37.01	—
Jing Jie	39.88	—
Gu Sui Bu	40.17	—
Yin Yang Huo	44.18	—
(Bai) Dou Kou	44.57	—
Wu Jia Pi	46.51	—
Hua Shi	46.60	—
Sang Bai Pi	46.77	—
Ma Bian Cao	47.34	—
He Huan Pi	47.59	—
E Zhu	49.08	—
Gui Zhi	49.21	—
Sang Ji Shend	49.72	—

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	24.79
62522	62522	0.00	24.78

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 2.1-2020-01

Validated substance/substance group **Wang Bu Liu Xing**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50390-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Wang Bu Liu Xing	3	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	G281HS057QM1	62181	40	from supplier
PhytoComm	Wang Bu Liu Xing	G281HS057QM1	62182	40	from supplier
PhytoComm	Wang Bu Liu Xing	G281H0456821	62463	40	from supplier
PhytoComm	Wang Bu Liu Xing	G281H0456821	62464	40	from supplier
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 39 821 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 *Wang Bu Liu Xing*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G281HS057QM1	62181 [†]	20
PhytoComm	Wang Bu Liu Xing	G281HS057QM1	62182 [†]	20
PhytoComm	Wang Bu Liu Xing	G281H0456821	62463 [†]	20
PhytoComm	Wang Bu Liu Xing	G281H0456821	62464 [†]	20
PhytoComm	Wang Bu Liu Xing	G281H0456921	62731 [†]	20
PhytoComm	Wang Bu Liu Xing	G281H0456921	62732 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	G281H0456521	1
PhytoComm	Wang Bu Liu Xing	G281H0456322	1
PhytoComm	Wang Bu Liu Xing	G281H456421	1
PhytoComm	Wang Bu Liu Xing	G281H0456221	1

- 849 spectra from 12 *Apo-Ident* customers from a total of 512 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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	0	4	849

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 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
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 *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
Fu Xiao Mai	6.90	—
Rou Gui	7.41	—
Zhu Ru	8.28	—
Ji Li	10.57	—
Fu Ling	11.74	—
Bai Bian Dou	11.83	—
Yi Yi Ren	12.16	—
Ling Zhi	13.02	—
Lian Zi	13.37	—
Sheng Jiang	13.68	—
He Huan Pi	14.36	—
Bai Shao Yao	14.92	—
Di Gu Pi	15.53	—
Bai Zi Ren	16.33	—
Gui Zhi	17.15	—
Ban Xia (Jiang)	17.69	—
Tao Ren	17.73	—
Ci Wu Jia	17.80	—
Tian Hua Fen	18.86	—
Shen Qu	18.89	—
Yu Jin	19.62	—
Huo Ma Ren	19.75	—
Shan Yao	21.06	—
Gua Lou	21.07	—
Lian Qiao	21.83	—
E Zhu	21.89	—
Tai Zi Shen	22.12	—
Zhe Bei Mu	22.33	—
Ren Dong Teng	22.76	—
Pi Pa Ye	22.88	—
Yuan Zhi	23.28	—
Qiang Huo	23.50	—
Bai Xian Pi	23.72	—
Gou Teng	23.73	—
Chen Pi	24.13	—
Ji Xue Teng	25.42	—
Chuan Xiong	26.35	—
Mo Yao	26.56	—
Fang Ji	26.76	—
Fu Zi	26.82	—
Mang Xiao	26.82	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tian Men Dong	27.08	—
Ye Jiao Teng	27.36	—
Mao Dong Qing	27.81	—
Wu Wei Zi	28.05	—
San Qi	28.47	—
Cang Er Zi	28.77	—
Che Qian Zi	29.25	—
Yan Hu Suo	29.31	—
Tu Fu Ling	29.50	—
Dan Shen	29.73	—
Ma Huang Gen	29.85	—
Du Zhong	29.89	—
Jiang Huang	29.91	—
Fo Shou	29.96	—
Da Zao	30.02	—
Yin Yang Huo	30.32	—
Sha Ren	30.64	—
Dang Shen	30.70	—
Shan Yu Rou	30.95	—
Long Yan Rou	31.08	—
Zi Su Zi	31.22	—
Chai Hu	31.42	—
Dan Dou Chi	32.27	—
Suan Zao Ren	32.90	—
Jiao Gu Lan	33.01	—
Wu Mei	33.29	—
Dang Gui Wei	33.40	—
Jin Yin Hua	34.37	—
Hou Po	34.38	—
Huang Lian	34.54	—
Zhi Mu	35.83	—
(Bai) Dou Kou	36.28	—
Fu Shen	36.59	—
Yi Mu Cao	36.75	—
Xi Xian Cao	36.80	—
Cang Zhu	37.45	—
Xin Yi	37.76	—
Sha Shen (Bei)	38.22	—
Sang Zhi	38.27	—
Dang Gui	38.95	—
Xiang Fu	39.45	—
Huang Qin	39.54	—
Chuan Lian Zi	39.58	—
Yu Zhu	39.92	—
She Gan	40.62	—
Gou Qi Zi	41.53	—
Xia Ku Cao	41.73	—
Pu Gong Ying	41.77	—
Zhi Ke	42.01	—
Hu Zhang	42.09	—
Qing Pi	42.24	—
Ze Xie	42.48	—
Nü Zhen Zi	42.54	—
Niu Bang Zi	42.80	—
Bai He	43.31	—
Wu Yao	44.33	—
Ba Ji Tian	44.72	—
Xie Bai	44.74	—
Mu Dan Pi	44.81	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Zhi Lian	45.96	—
Ma Huang	47.37	—
Jie Geng	47.43	—
Du Huo	47.89	—
Zi Hua Di Ding	48.45	—
Gan Jiang	49.08	—
Ma Bian Cao	49.14	—
Lai Fu Zi	49.94	—

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	27.03
62182	62182	0.00	6.90
62463	62463	0.00	22.34
62464	62464	0.00	21.89
62732	62732	0.00	26.82
62181	62181	0.00	7.78

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 2.1-2020-01

Validated substance/substance group **Wei Ling Xian**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60543-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Wei Ling Xian; Clematidis radix

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Wei Ling Xian	1	0	3

Second-stage model

For differentiation of the substance/substance group *Wei Ling Xian* 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 *Wei Ling Xian*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Wei Ling Xian	G075H0943822	62381	40	from supplier
PhytoComm	Wei Ling Xian	G075H0943822	62382	40	from supplier

Validation samples

A total of 39 821 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 *Wei Ling Xian*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *Wei Ling Xian*.
- 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	Wei Ling Xian	G075H0943822	62381 [†]	20
PhytoComm	Wei Ling Xian	G075H0943822	62382 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 4 batches from the substance/substance group *Wei Ling Xian*.
- 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	Wei Ling Xian	G075H0943223	1
Phytocomm	Wei Ling Xian	G075H0943422	1
Phytocomm	Wei Ling Xian	g075h0943223	1
PhytoComm	Wei Ling Xian	G075H0943223	2
PhytoComm	Wei Ling Xian	G075H0943223	1

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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 *Wei Ling Xian* 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 *Wei Ling Xian* 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	25 721
Type B	0	40	0	13 127
Type C	0	0	6	847

The substance/substance group *Wei Ling Xian* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8235 %)	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 *Wei Ling Xian* 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	14.15	—
Ku Shen	15.73	—
Gua Lou	17.15	—
Huang Bo	18.56	—
Gui Zhi	18.74	—
Xu Duan	20.52	—
Cang Er Zi	21.13	—
Ze Xie	24.03	—
Nü Zhen Zi	24.55	—
Lian Qiao	25.09	—
Qing Pi	25.18	—
She Gan	25.83	—
Xie Bai	26.16	—
(Shi) Chang Pu	26.92	—
Qiang Huo	27.38	—
Zhi Zi	27.50	—
Wu Yao	27.70	—
Fu Zi	27.73	—
Xiang Fu	30.54	—
Chai Hu	31.48	—
Zi Wan	31.63	—
Jiang Huang	32.02	—
Huo Ma Ren	32.32	—
Niu Bang Zi	32.59	—
Cang Zhu	33.92	—
Ju Hua	34.01	—
Sha Ren	34.09	—
Ma Bian Cao	34.40	—
Wu Zhu Yu	34.57	—
Fang Feng	34.93	—
Hou Po	35.42	—
(Huai) Niu Xi	35.83	—
Shan Yao	35.85	—
Di Gu Pi	35.94	—
Gou Qi Zi	35.96	—
Gan Jiang	36.01	—
Shen Qu	36.03	—
Du Huo	37.08	—
(Bai) Dou Kou	37.47	—
Bu Gu Zhi	37.54	—
Pu Gong Ying	37.64	—
Sang Zhi	37.65	—
Mang Xiao	37.78	—
Gan Cao	38.21	—
Wu Wei Zi	39.41	—
Dan Shen	39.64	—
Yin Yang Huo	39.73	—
He Huan Pi	40.14	—
Chen Pi	40.27	—
Dang Gui	40.59	—
Xin Yi	40.64	—
Ji Li	41.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Hu Zhang	42.31	—
Gu Sui Bu	42.33	—
Jin Yin Hua	44.03	—
Dang Gui Wei	44.16	—
Hong Jing Tian	44.19	—
Dong Gua Zi	44.42	—
Zhi Mu	44.73	—
Bai Mao Gen	46.25	—
Chi Shao (Yao)	46.28	—
Huang Qi	47.22	—
Xi Xian Cao	47.27	—
Chuan Xiong	47.51	—
Xian He Cao	47.54	—
Zi Su Zi	48.09	—
Yin Chen Hao	48.31	—
Tu Fu Ling	48.75	—
Sang Bai Pi	49.12	—
Jie Geng	49.66	—
Rou Cong Rong	49.69	—
Tian Hua Fen	49.89	—
Huang Lian	50.03	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Wei Ling Xian* 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
62382	62382	0.00	14.31
62381	62381	0.00	14.15

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 2.1-2020-01

Validated substance/substance group **Wu Jia Pi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002259-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Wu Jia Pi; Acanthopanax cortex

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 848 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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
Sha Ren	14.23	—
(Sheng) Di Huang	16.19	—
Yin Chen Hao	17.02	—
Sang Bai Pi	18.43	—
Bu Gu Zhi	19.62	—
Jing Jie	21.07	—
Yu Xing Cao	21.84	—
Dan Zhu Ye	22.60	—
Du Zhong	26.06	—
Pu Gong Ying	29.49	—
Xuan Shen	31.42	—
Jin Qian Cao	35.36	—
Sang Ji Shend	35.51	—
Nü Zhen Zi	35.80	—
Ge Gen	37.39	—
Xu Duan	37.45	—
Xian He Cao	39.09	—
Yin Yang Huo	39.52	—
Ma Bian Cao	40.48	—
Qing Hao	43.18	—
He Shou Wu	43.57	—
Gu Sui Bu	44.72	—
Hua Shi	47.19	—

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
62758	62758	0.00	14.23
62757	62757	0.00	15.00

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 2.1-2020-01

Validated substance/substance group **Wu Mei**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60021-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	25 641
Type B	0	80	0	13 087
Type C	0	0	2	851

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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
Wu Wei Zi	7.61	—
Qing Pi	17.95	—
Sang Ye	18.73	—
Shan Yao	20.00	—
Cang Er Zi	20.16	—
Yi Mu Cao	20.67	—
Gui Zhi	20.93	—
Huang Bai	23.41	—
Qiang Huo	25.72	—
Fu Zi	27.28	—
Xi Xian Cao	27.49	—
Huang Lian	29.43	—
Hong Jing Tian	29.73	—
Du Zhong	30.29	—
Mao Dong Qing	30.32	—
Jiang Huang	30.52	—
Shen Qu	30.90	—
Hou Po	30.96	—
Jin Yin Hua	31.15	—
Xin Yi	31.20	—
Fang Ji	31.91	—
Xia Ku Cao	31.92	—
Jiao Gu Lan	32.06	—
Ma Huang	34.37	—
Hong Hua	34.40	—
Zi Hua Di Ding	35.50	—
Pu Gong Ying	36.16	—
Yin Yang Huo	36.17	—
Zhi Zi	36.38	—
Guang Huo Xiang	37.60	—
Tu Fu Ling	38.57	—
Lian Qiao	38.87	—
Wu Yao	39.02	—
Bu Gu Zhi	39.26	—
Bai Zhu	39.94	—
Dan Shen	40.63	—
Chai Hu	40.68	—
Zhi Ke	41.00	—
Hu Zhang	41.36	—
Ye Jiao Teng	41.75	—
He Huan Pi	41.99	—
Gan Cao	42.11	—
Wei Ling Xian	44.80	—
Shan Yu Rou	46.18	—
Hua Shi	46.32	—
Qing Hao	46.39	—
Nü Zhen Zi	46.40	—
Wang Bu Liu Xing	46.79	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gua Lou	46.81	—
Yu Xing Cao	47.09	—
Di Gu Pi	47.64	—
Bai Mao Gen	47.82	—
Gou Teng	48.14	—
Mang Xiao	48.57	—
Ai Ye	49.81	—
Sha Ren	49.92	—

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
62478	62478	0.00	31.36
62715	62715	0.00	7.61
62716	62716	0.00	8.54
62477	62477	0.00	32.42

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 2.1-2020-01

Validated substance/substance group **Wu Wei Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60001-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G220HS048RP1	62265	40	from supplier
PhytoComm	Wu Wei Zi	G220HS048RP1	62266	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G220HS048RP1	62265 [†]	20
PhytoComm	Wu Wei Zi	G220HS048RP1	62266 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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

- 850 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 *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	25 561
Type B	0	120	0	13 047
Type C	0	0	3	850

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8257 %)	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	10.40	—
Jiao Gu Lan	12.69	—
Qing Pi	14.28	—
Zhi Ke	15.00	—
Shan Yao	15.30	—
Lian Qiao	16.14	—
Huang Bai	16.21	—
Wu Mei	16.26	—
Fang Ji	17.07	—
Suan Zao Ren	17.23	—
Cang Er Zi	18.16	—
Sang Ye	20.54	—
Yan Hu Suo	20.67	—
Cang Zhu	21.06	—
Shan Yu Rou	21.58	—
San Qi	22.87	—
Mu Dan Pi	23.43	—
Gui Zhi	23.69	—
Zhe Bei Mu	24.28	—
Chuan Xiong	24.35	—
Dan Shen	26.01	—
Hou Po	26.13	—
Yi Mu Cao	26.65	—
Shen Qu	27.10	—
Fu Zi	27.45	—
Du Zhong	27.63	—
Zi Hua Di Ding	27.73	—
Qiang Huo	27.85	—
Bai Zi Ren	28.21	—
Bai Bian Dou	28.51	—
He Huan Pi	28.54	—
Ye Jiao Teng	29.77	—
Mao Dong Qing	29.83	—
Bai Zhu	29.99	—
Zhi Mu	30.04	—
Gou Qi Zi	30.27	—
Hong Hua	30.45	—
Dang Shen	30.53	—
Di Gu Pi	30.87	—
Ji Li	30.97	—
Bai Shao Yao	31.56	—
Pi Pa Ye	31.60	—
Huang Lian	31.60	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yin Yang Huo	31.75	—
Tai Zi Shen	31.81	—
Hong Jing Tian	31.82	—
Ban Xia (Jiang)	31.83	—
Chai Hu	31.90	—
Gua Lou	32.35	—
Xi Xian Cao	32.82	—
Yu Zhu	33.00	—
Ma Huang	33.13	—
Yuan Zhi	33.36	—
Lian Zi	33.54	—
Chen Pi	33.55	—
Tao Ren	33.83	—
Dang Gui	35.42	—
Ren Dong Teng	35.85	—
Xia Ku Cao	36.57	—
Huang Qin	37.23	—
Chi Shao (Yao)	37.38	—
Huo Ma Ren	37.60	—
Gan Cao	38.07	—
Ling Zhi	38.75	—
Jiang Huang	38.80	—
Chuang Mu Xiang	38.93	—
Bai Xian Pi	39.67	—
Sheng Jiang	40.78	—
Xin Yi	41.24	—
Yi Yi Ren	42.83	—
Che Qian Zi	42.83	—
Wu Yao	43.14	—
Bai Mao Gen	43.18	—
Ji Xue Teng	43.38	—
Tu Fu Ling	44.45	—
Mang Xiao	44.65	—
Dang Gui Wei	44.76	—
Bu Gu Zhi	45.02	—
Zhu Ru	45.79	—
Pu Gong Ying	45.93	—
Zhi Zi	46.01	—
Sang Zhi	46.16	—
Dan Dou Chi	46.25	—
Tian Hua Fen	47.86	—
Fu Ling	48.39	—
Rou Gui	48.45	—
Fu Xiao Mai	48.57	—
Gou Teng	49.01	—
Zi Wan	49.32	—
Dong Kui Zi	49.49	—
(Shi) Chang Pu	49.84	—

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	11.68
62552	62552	0.00	10.40
62265	62265	0.00	14.46
62706	62706	0.00	18.22
62266	62266	0.00	14.28
62705	62705	0.00	18.16

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 2.1-2020-01

Validated substance/substance group **Wu Yao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60040-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Wu Yao; Linderæ radix

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Wu Yao	2	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	G140H1025721	62233	40	from supplier
PhytoComm	Wu Yao	G140H1025721	62234	40	from supplier
PhytoComm	Wu Yao	G140H1025921	62741	40	from supplier
PhytoComm	Wu Yao	G140H1025921	62742	40	from supplier

Validation samples

A total of 39 821 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 Yao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G140H1025721	62233 [†]	20
PhytoComm	Wu Yao	G140H1025721	62234 [†]	20
PhytoComm	Wu Yao	G140H1025921	62741 [†]	20
PhytoComm	Wu Yao	G140H1025921	62742 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G140H1025521	2
PhytoComm	Wu Yao	G140H1025221	2
PhytoComm	Wu Yao	G140H1025521	1

- 847 spectra from 12 *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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	9	3	3	838

The substance/substance group *Wu 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	100.0000 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.2248 % (> 98.6366 %)	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 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
Chen Pi	13.52	—
Sha Ren	13.81	—
Yu Jin	14.08	—
Gui Zhi	14.12	—
He Huan Pi	15.19	—
Yi Mu Cao	15.41	—
Lu Gen	15.65	—
Xian He Cao	15.86	—
Gua Lou	16.29	—
Ji Li	16.51	—
Jiang Huang	16.67	—
Gan Jiang	17.34	—
Fang Feng	18.11	—
Hou Po	19.40	—
Fu Zi	19.79	—
Mu Dan Pi	19.80	—
Xu Duan	19.83	—
Ma Bian Cao	21.14	—
Dan Shen	21.46	—
Tu Fu Ling	21.69	—
Gou Qi Zi	21.69	—
Gu Sui Bu	21.89	—
Sang Zhi	21.90	—
(Shi) Chang Pu	22.54	—
Chai Hu	22.64	—
(Bai) Dou Kou	22.95	—
Huang Jing	23.63	—
He Shou Wu	23.65	—
Xiang Fu	23.76	—
Qiang Huo	24.39	—
Chuan Xiong	25.03	—
Du Huo	25.16	—
Mang Xiao	25.25	—
Zi Su Zi	25.37	—
Ju Hua	25.57	—
Hong Jing Tian	25.76	—
Yin Yang Huo	26.05	—
She Gan	26.48	—
Lian Qiao	26.53	—
Yin Chen Hao	26.98	—
Sang Bai Pi	27.30	—
Ba Ji Tian	27.50	—
Yan Hu Suo	27.59	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Dang Gui	27.72	—
Zi Wan	27.86	—
Qing Pi	28.50	—
Sang Ji Shend	28.88	—
Wu Zhu Yu	29.14	—
Huang Bo	29.87	—
Nü Zhen Zi	30.41	—
Chuan Lian Zi	30.64	—
Xie Bai	30.70	—
Huo Ma Ren	31.12	—
E Zhu	31.58	—
Dang Gui Wei	32.31	—
Shan Yao	32.78	—
Di Gu Pi	32.88	—
Hu Zhang	33.07	—
Bai Mao Gen	33.11	—
Zhi Mu	33.85	—
Cang Zhu	34.59	—
San Leng	35.34	—
Bu Gu Zhi	35.45	—
Long Dan (Cao)	35.69	—
Xin Yi	35.84	—
Ge Gen	36.17	—
(Huai) Niu Xi	36.72	—
Huang Lian	36.87	—
Niu Bang Zi	36.98	—
Sheng Jiang	37.27	—
Zhi Ke	37.48	—
Huang Qi	37.52	—
Cang Er Zi	38.07	—
Tian Hua Fen	39.01	—
Ze Xie	39.78	—
Dang Shen	40.61	—
Ku Shen	40.64	—
Bai Bian Dou	41.00	—
Lai Fu Zi	42.14	—
Dan Zhu Ye	42.62	—
Wei Ling Xian	43.02	—
Pu Gong Ying	45.08	—
Xia Ku Cao	45.99	—
Jie Geng	46.16	—
Zhi Shi	46.67	—
Du Zhong	46.74	—
Sang Ye	47.02	—
Mao Dong Qing	47.25	—
Wang Bu Liu Xing	47.70	—
San Qi	48.23	—
Wu Wei Zi	48.26	—
Xi Xian Cao	48.44	—

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
62233	62233	0.00	14.17
62742	62742	0.00	13.52
62234	62234	0.00	13.81
62741	62741	0.00	13.80

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 2.1-2020-01

Validated substance/substance group **Wu Zhu Yu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002260-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	25 721
Type B	0	40	0	13 127
Type C	0	1	1	851

The substance/substance group *Wu Zhu Yu* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8279 %)	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
Xian He Cao	21.05	—
Huang Bo	21.30	—
Nü Zhen Zi	21.31	—
Ju Hua	22.50	—
Qiang Huo	22.83	—
Xu Duan	23.59	—
Hou Po	26.05	—
Ma Bian Cao	26.65	—
Yin Chen Hao	26.74	—
Wu Yao	26.89	—
Gu Sui Bu	27.03	—
Dan Shen	27.87	—
Ku Shen	27.87	—
Wei Ling Xian	28.06	—
Yi Mu Cao	28.72	—
Dan Zhu Ye	30.18	—
Sha Ren	31.92	—
Sang Ji Shend	32.17	—
Fang Feng	32.52	—
Ze Xie	32.70	—
Hong Jing Tian	33.94	—
Fu Zi	35.54	—
Gui Zhi	35.69	—
Lian Qiao	36.15	—
Yu Jin	37.54	—
Xie Bai	37.58	—
He Shou Wu	38.12	—
Cang Er Zi	38.42	—
Zi Wan	38.85	—
Qing Pi	40.53	—
He Huan Pi	41.07	—
Xiang Fu	41.43	—
Yin Yang Huo	41.59	—
Huang Lian	41.98	—
Mang Xiao	42.05	—
She Gan	42.18	—
Jing Jie	42.92	—
Chen Pi	42.93	—
Jiang Huang	43.07	—
Gou Qi Zi	43.17	—
Chai Hu	44.16	—
(Shi) Chang Pu	45.63	—
Bu Gu Zhi	46.18	—
Sang Bai Pi	46.41	—
Cang Zhu	46.63	—
Zhi Shi	47.15	—
Niu Bang Zi	47.49	—
Ge Gen	47.61	—
Gua Lou	48.41	—
Huang Jing	48.76	—
Rou Cong Rong	49.85	—

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
62782	62782	0.00	21.05
62781	62781	0.00	21.30

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 2.1-2020-01

Validated substance/substance group **Xi Xian Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60103-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Xi Xian Cao	1	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

Validation samples

A total of 39 821 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 *Xi Xian Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 *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

- 13 127 spectra from a total of 319 batches from further 172 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

- 850 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *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
Tu Fu Ling	8.26	—
Xin Yi	9.18	—
Pu Gong Ying	14.08	—
Xia Ku Cao	15.35	—
Nü Zhen Zi	15.39	—
Shan Yao	19.59	—
Du Zhong	19.85	—
Sang Ye	19.90	—
Yu Jin	20.09	—
Huang Lian	20.18	—
Fang Ji	23.80	—
Jiang Huang	24.30	—
Zhi Ke	24.58	—
Gui Zhi	24.71	—
Hu Zhang	25.13	—
Jiao Gu Lan	26.68	—
Qiang Huo	27.74	—
Bai Xian Pi	27.91	—
Mao Dong Qing	28.32	—
Ling Zhi	28.46	—
Cang Er Zi	28.64	—
Dan Shen	29.15	—
Han Lian Cao	30.06	—
Hou Po	30.29	—
Bai Hua She She Cao	30.43	—
Wu Mei	30.89	—
Fu Zi	31.48	—
Qing Pi	32.31	—
Yin Yang Huo	32.34	—
Yan Hu Suo	32.37	—
Ye Jiao Teng	32.89	—
Wang Bu Liu Xing	34.78	—
Pi Pa Ye	36.12	—
Di Gu Pi	36.84	—
Che Qian Zi	37.74	—
He Huan Pi	37.91	—
Wu Wei Zi	37.98	—
Dan Dou Chi	38.19	—
Mang Xiao	38.50	—
Jin Yin Hua	38.72	—
Ji Xue Teng	39.18	—
Yi Mu Cao	39.44	—
Sha Ren	39.96	—
Hong Jing Tian	40.31	—
Lian Zi	40.67	—
E Zhu	41.26	—
Ma Huang	41.95	—
Shen Qu	42.25	—
Zi Hua Di Ding	42.42	—
Ji Li	42.44	—
Bai Bian Dou	43.68	—
Xiang Fu	44.65	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ren Dong Teng	44.69	—
Wei Ling Xian	44.78	—
Bu Gu Zhi	44.86	—
Wu Yao	45.10	—
Ze Xie	45.14	—
Bai Shao Yao	45.47	—
Chuan Xiong	46.43	—
Chai Hu	46.82	—
Mu Dan Pi	47.09	—
Lian Qiao	47.70	—
Lu Gen	49.11	—
Hong Hua	49.45	—
Gou Qi Zi	49.46	—
Huang Bai	49.94	—

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
62556	62556	0.00	8.67
62555	62555	0.00	8.26

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 2.1-2020-01

Validated substance/substance group **Xia Ku Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60080-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Xia Ku Cao	2	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	G202HS232QM1	62191	40	from supplier
PhytoComm	Xia Ku Cao	G202HS232QM1	62192	40	from supplier
PhytoComm	Xia Ku Cao	G202HS232RM1	62283	40	from supplier
PhytoComm	Xia Ku Cao	G202HS232RM1	62284	40	from supplier

Validation samples

A total of 39 821 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 *Xia Ku Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G202HS232QM1	62191 [†]	20
PhytoComm	Xia Ku Cao	G202HS232QM1	62192 [†]	20
PhytoComm	Xia Ku Cao	G202HS232RM1	62283 [†]	20
PhytoComm	Xia Ku Cao	G202HS232RM1	62284 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G202H1064221	2
Phytocomm	Xia Ku Cao	g202h1064121	1
Phytocomm	Xia Ku Cao	G202H1064421	1
PhytoComm	Xia Ku Cao	G202H1064221	2

- 847 spectra from 12 *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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	6	847

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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8235 %)	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 *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	7.38	—
Bai Hua She She Cao	11.58	—
Xi Xian Cao	13.83	—
Hu Zhang	16.68	—
Huang Lian	19.02	—
Xin Yi	19.87	—
Tu Fu Ling	20.43	—
Han Lian Cao	20.62	—
Jiang Huang	22.19	—
Xuan Shen	23.21	—
Gui Zhi	24.22	—
Pu Gong Ying	24.60	—
Shan Yao	25.58	—
Sang Ye	26.58	—
Yin Yang Huo	31.43	—
Nü Zhen Zi	31.48	—
Yin Chen Hao	31.63	—
Yi Mu Cao	32.85	—
He Huan Pi	33.15	—
Sha Ren	33.52	—
Huang Bo	34.05	—
Fu Zi	35.33	—
Bu Gu Zhi	36.28	—
Cang Er Zi	37.11	—
Dan Shen	38.08	—
He Shou Wu	38.86	—
Du Zhong	39.16	—
Qiang Huo	39.47	—
Gu Sui Bu	40.19	—
Mang Xiao	40.81	—
Dan Zhu Ye	41.18	—
Wu Zhu Yu	41.62	—
E Zhu	42.61	—
Yu Jin	42.89	—
Jiao Gu Lan	43.92	—
Sang Bai Pi	45.01	—
Bai Xian Pi	45.11	—
Shen Qu	45.74	—
Sang Ji Shend	47.02	—
Xiang Fu	47.09	—
Yan Hu Suo	47.19	—
Zhi Ke	47.46	—
Xian He Cao	47.85	—
Ma Bian Cao	47.91	—
Rou Cong Rong	48.11	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Gou Teng	48.30	—
Ji Li	48.52	—
Xu Duan	48.88	—
Hou Po	49.05	—
Qing Pi	49.21	—
Jing Jie	49.33	—
Wu Mei	49.52	—

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
62192	62192	0.00	11.58
62191	62191	0.00	11.73
62283	62283	0.00	7.38
62284	62284	0.00	7.39

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 2.1-2020-01

Validated substance/substance group **Xian He Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60591-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Xian He Cao; Agrimoniae herba

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Xian He Cao	1	0	2

Second-stage model

For differentiation of the substance/substance group *Xian He 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 *Xian He Cao*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Xian He Cao	G008H0547921	62585	40	from supplier
PhytoComm	Xian He Cao	G008H0547921	62586	40	from supplier

Validation samples

A total of 39 821 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 He Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 He 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	Xian He Cao	G008H0547921	62585 [†]	20
PhytoComm	Xian He Cao	G008H0547921	62586 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Xian He 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	Xian He Cao	g008h0547321	1
Phytocomm	Xian He Cao	G008H0547321	1
Phytocomm	Xian He Cao	G008H0547121	1

- 850 spectra from 12 *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 *Xian He 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 *Xian He 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	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

The substance/substance group *Xian He 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 He 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
Gu Sui Bu	13.61	—
Yu Jin	14.30	—
He Huan Pi	16.24	—
Sang Ji Shend	16.87	—
He Shou Wu	17.38	—
Yi Mu Cao	19.54	—
Hong Jing Tian	19.72	—
Sha Ren	21.23	—
Wu Yao	21.55	—
Chen Pi	21.56	—
Ma Bian Cao	21.69	—
Dan Shen	23.12	—
Hou Po	23.60	—
Gui Zhi	23.97	—
Sang Bai Pi	24.15	—
Xu Duan	24.32	—
Fu Zi	25.55	—
Yin Chen Hao	25.88	—
Yin Yang Huo	26.36	—
Fang Feng	26.43	—
Wu Zhu Yu	27.35	—
Qiang Huo	28.88	—
Lu Gen	28.99	—
Nü Zhen Zi	29.89	—
Ji Li	29.93	—
Ge Gen	30.59	—
Huang Bo	31.23	—
Jiang Huang	32.59	—
Dan Zhu Ye	32.69	—
Bu Gu Zhi	32.88	—
Mang Xiao	32.90	—
Huang Jing	38.55	—
Xiang Fu	39.27	—
Lian Qiao	39.55	—
(Bai) Dou Kou	39.90	—
Hu Zhang	40.38	—
Zhi Shi	41.35	—
Ju Hua	43.15	—
(Shi) Chang Pu	43.76	—
E Zhu	43.77	—
Gou Qi Zi	45.33	—
She Gan	47.67	—
Shan Yao	48.31	—
Ku Shen	48.35	—
Chai Hu	48.62	—
Sang Ye	48.98	—
Chuan Xiong	49.62	—
Qing Pi	49.68	—
Hua Shi	50.46	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the sub-

stance/substance group *Xian He 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
62586	62586	0.00	13.61
62585	62585	0.00	15.77

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 2.1-2020-01

Validated substance/substance group **Xiang Fu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60185-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G088H0930821	62333	40	from supplier
PhytoComm	Xiang Fu	G088H0930821	62334	40	from supplier
PhytoComm	Xiang Fu	G088H0930921	62777	40	from supplier
PhytoComm	Xiang Fu	G088H0930921	62778	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G088H0930821	62333 [†]	20
PhytoComm	Xiang Fu	G088H0930821	62334 [†]	20
PhytoComm	Xiang Fu	G088H0930921	62777 [†]	20
PhytoComm	Xiang Fu	G088H0930921	62778 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G088H0930521	5
PhytoComm	Xiang Fu	G088H0930321	1
PhytoComm	Xiang Fu	G088H0930521	1

- 845 spectra from 12 *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 *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	25 641
Type B	0	80	0	13 087
Type C	0	4	4	845

The substance/substance group *Xiang Fu* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8230 %)	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 *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
(Shi) Chang Pu	8.84	—
Ze Xie	10.10	—
Gui Zhi	11.26	—
Qing Pi	11.84	—
Sang Zhi	12.10	—
She Gan	12.17	—
Zi Wan	13.93	—
Gou Qi Zi	14.25	—
Yi Mu Cao	14.39	—
Du Huo	14.66	—
Lu Gen	15.34	—
Chai Hu	15.92	—
Dang Gui Wei	16.27	—
Gua Lou	16.40	—
Lian Qiao	17.95	—
Dang Gui	18.56	—
Cang Zhu	18.67	—
Di Gu Pi	19.60	—
Gan Jiang	19.81	—
Wu Yao	20.02	—
Cang Er Zi	20.41	—
Bai Mao Gen	20.44	—
Chuan Xiong	21.44	—
Niu Bang Zi	22.16	—
Ku Shen	22.20	—
Wei Ling Xian	22.89	—
Jiang Huang	22.93	—
Tian Hua Fen	23.51	—
Ji Li	23.71	—
San Leng	23.78	—
(Huai) Niu Xi	24.05	—
Fang Feng	24.53	—
Shan Yao	24.71	—
Huang Qi	25.52	—
Zhi Mu	25.82	—
Mu Dan Pi	25.90	—
Huo Ma Ren	26.04	—
Zhi Ke	26.23	—
Chi Shao (Yao)	26.31	—
Pu Gong Ying	26.57	—
Jie Geng	27.11	—
Xie Bai	27.90	—
(Bai) Dou Kou	28.01	—
E Zhu	28.44	—
Wu Wei Zi	28.50	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ju Hua	28.91	—
Hou Po	29.13	—
Yan Hu Suo	29.35	—
Hong Jing Tian	29.64	—
Fu Zi	29.67	—
Sha Ren	29.84	—
Yuan Zhi	30.11	—
Huang Lian	30.12	—
Chen Pi	30.33	—
Mang Xiao	30.43	—
Zi Su Zi	30.63	—
Chuan Lian Zi	30.71	—
Xu Duan	30.87	—
Xin Yi	31.26	—
Tu Fu Ling	31.69	—
Zhi Zi	32.63	—
Jin Yin Hua	32.95	—
Ba Ji Tian	33.11	—
Huang Bo	33.62	—
Long Dan (Cao)	34.06	—
Dang Shen	34.60	—
Sha Shen (Bei)	35.11	—
He Huan Pi	35.66	—
Xian He Cao	35.77	—
Gan Cao	37.24	—
Shan Yu Rou	37.50	—
Bai Shao Yao	37.58	—
Sang Ye	37.75	—
Mao Dong Qing	38.13	—
Lai Fu Zi	38.18	—
Qiang Huo	38.33	—
Dong Gua Zi	38.95	—
Sheng Jiang	39.00	—
Zi Hua Di Ding	39.52	—
Dan Shen	39.90	—
Xi Xian Cao	39.91	—
Fang Ji	40.29	—
Du Zhong	40.70	—
Nü Zhen Zi	40.98	—
Bai Zhu	41.04	—
Lian Zi	41.27	—
Huang Bai	41.47	—
Yu Jin	42.22	—
Hu Zhang	42.76	—
Mu Gua	42.89	—
Dan Dou Chi	42.92	—
Jiao Gu Lan	43.44	—
Ma Bian Cao	43.88	—
Gu Sui Bu	44.15	—
Bai Xian Pi	44.18	—
Ye Jiao Teng	44.61	—
Yin Yang Huo	45.14	—
Ji Xue Teng	45.22	—
Bai He	45.42	—
Wu Zhu Yu	45.45	—
Qin Jiao	45.64	—
Bu Gu Zhi	45.87	—
Che Qian Zi	46.17	—
Ge Gen	46.86	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ban Zhi Lian	48.55	—
Wu Mei	48.57	—
Ren Dong Teng	48.63	—
Sang Ji Shend	49.01	—
Shen Qu	49.13	—
Sang Bai Pi	49.78	—
He Shou Wu	49.88	—
Gou Teng	50.08	—

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
62334	62334	0.00	8.84
62777	62777	0.00	10.00
62333	62333	0.00	8.99
62778	62778	0.00	9.82

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 2.1-2020-01

Validated substance/substance group **Xie Bai**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60362-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G504H1830821	62321	40	from supplier
PhytoComm	Xie Bai	G504H1830821	62322	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G504H1830821	62321 [†]	20
PhytoComm	Xie Bai	G504H1830821	62322 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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	G504H1830421	1
Phytocomm	Xie Bai	G504H1830422	1
PhytoComm	Xie Bai	G504H1830321	1

- 850 spectra from 12 *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 *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	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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
Gua Lou	12.41	—
Huo Ma Ren	13.87	—
Gan Jiang	15.71	—
Zi Su Zi	16.02	—
She Gan	16.24	—
Zhi Mu	17.07	—
(Huai) Niu Xi	19.34	—
Gou Qi Zi	19.66	—
Cang Zhu	20.33	—
Chai Hu	20.45	—
Cang Er Zi	23.92	—
Fang Feng	24.01	—
Bai Bian Dou	24.14	—
Yi Mu Cao	24.32	—
Sang Zhi	24.40	—
Xiang Fu	25.38	—
(Shi) Chang Pu	25.58	—
Jie Geng	25.70	—
Ji Li	26.43	—
Di Gu Pi	26.83	—
Gui Zhi	26.86	—
Yan Hu Suo	27.20	—
Mang Xiao	27.37	—
Niu Bang Zi	27.77	—
Qing Pi	28.50	—
Jiang Huang	28.50	—
Huang Qi	28.95	—
Shan Yao	29.07	—
Dang Shen	29.43	—
Wei Ling Xian	29.59	—
Du Huo	29.66	—
Hou Po	30.93	—
Lian Qiao	31.11	—
Dong Gua Zi	31.30	—
Ku Shen	31.61	—
Dang Gui	31.72	—
Zi Wan	32.04	—
Sha Ren	32.06	—
Qiang Huo	32.09	—
Chen Pi	32.11	—
Ba Ji Tian	32.38	—
(Bai) Dou Kou	32.43	—
Dang Gui Wei	32.48	—
Tian Hua Fen	33.13	—
San Qi	33.23	—
Wu Yao	33.69	—
Huang Bo	33.81	—
Mu Gua	34.36	—
Sheng Jiang	34.39	—
Chuan Xiong	34.63	—
Bai Mao Gen	35.26	—
Nü Zhen Zi	35.50	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ju Hua	35.71	—
E Zhu	35.73	—
Wang Bu Liu Xing	35.81	—
Xu Duan	36.54	—
Lai Fu Zi	36.78	—
Bai Zhu	37.16	—
Zhi Zi	37.39	—
He Huan Pi	37.96	—
Chuan Lian Zi	38.09	—
Ze Xie	39.52	—
San Leng	40.06	—
Bai He	40.27	—
Ma Bian Cao	40.69	—
Yuan Zhi	40.73	—
Chi Shao (Yao)	40.75	—
Long Dan (Cao)	40.81	—
Lu Gen	40.98	—
Xian He Cao	41.59	—
Jin Yin Hua	41.60	—
Wu Wei Zi	41.73	—
Xin Yi	41.78	—
Yin Yang Huo	42.13	—
Zhi Ke	42.59	—
Mu Dan Pi	43.20	—
Huang Lian	43.25	—
Lian Zi	44.15	—
Tu Fu Ling	44.63	—
Hong Jing Tian	45.44	—
Ban Xia (Jiang)	45.55	—
Sha Shen (Bei)	46.24	—
Pu Gong Ying	46.62	—
Dan Shen	46.77	—
Fu Zi	46.86	—
Bai Xian Pi	47.14	—
Yu Jin	47.15	—
Sang Bai Pi	47.33	—
Bu Gu Zhi	49.29	—
Huang Jing	49.34	—
Fang Ji	50.20	—

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
62322	62322	0.00	12.41
62321	62321	0.00	12.67

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 2.1-2020-01

Validated substance/substance group **Xin Yi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60231-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G153H0758821	62351	40	from supplier
PhytoComm	Xin Yi	G153H0758821	62352	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G153H0758821	62351 [†]	20
PhytoComm	Xin Yi	G153H0758821	62352 [†]	20
PhytoComm	Xin Yi	G153HS139SH1	62603 [†]	20
PhytoComm	Xin Yi	G153HS139SH1	62604 [†]	20
PhytoComm	Xin Yi	G153HS139SP1	62725 [†]	20
PhytoComm	Xin Yi	G153HS139SP1	62726 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	G153H0758521	1
PhytoComm	Xin Yi	G153H0758321	1

- 851 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 *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	25 561
Type B	0	120	0	13 047
Type C	0	0	2	851

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *Ma-halanobis 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	6.35	—
Pu Gong Ying	7.31	—
Tu Fu Ling	10.55	—
Xia Ku Cao	12.85	—
Nü Zhen Zi	13.14	—
Bai Hua She She Cao	14.66	—
Sang Ye	15.78	—
Han Lian Cao	18.59	—
Shan Yao	18.60	—
Zhi Shi	19.47	—
Hu Zhang	20.25	—
Xuan Shen	20.73	—
Huang Lian	20.94	—
Jiang Huang	21.26	—
Qing Hao	22.97	—
Rou Cong Rong	24.82	—
Gui Zhi	25.12	—
Yin Yang Huo	25.31	—
Yi Mu Cao	25.64	—
Du Zhong	25.97	—
Yu Jin	26.42	—
Cang Er Zi	26.63	—
Fang Ji	30.66	—
Fu Zi	30.96	—
Huang Bo	31.07	—
Yan Hu Suo	31.83	—
Dan Shen	31.89	—
Qiang Huo	31.98	—
Zhi Ke	33.32	—
Bu Gu Zhi	33.92	—
Mao Dong Qing	34.85	—
He Huan Pi	35.25	—
Bai Xian Pi	35.38	—
Yin Chen Hao	35.55	—
Gou Teng	35.62	—
Sha Ren	35.64	—
Hou Po	35.64	—
Jiao Gu Lan	36.10	—
Wu Mei	36.91	—
Ling Zhi	37.07	—
Qing Pi	37.14	—
Wang Bu Liu Xing	37.66	—
Dan Zhu Ye	37.72	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ye Jiao Teng	38.39	—
Shen Qu	38.94	—
Wu Wei Zi	39.67	—
Sang Ji Shend	40.42	—
Mang Xiao	40.45	—
Di Gu Pi	41.35	—
Pi Pa Ye	42.20	—
Ji Li	42.29	—
Wei Ling Xian	42.48	—
E Zhu	43.08	—
Hong Jing Tian	43.25	—
Xiang Fu	44.16	—
Wu Yao	44.65	—
Che Qian Zi	44.77	—
Jin Yin Hua	45.08	—
Dan Dou Chi	45.20	—
Jing Jie	45.26	—
Gu Sui Bu	45.32	—
Guang Huo Xiang	45.50	—
Ze Xie	45.79	—
Xu Duan	46.95	—
He Shou Wu	47.27	—
Wu Zhu Yu	47.28	—
Gou Qi Zi	47.79	—
Ma Huang	48.06	—
Sang Bai Pi	48.12	—
Lu Gen	48.20	—
Jin Qian Cao	48.48	—
Hua Shi	48.64	—
Chai Hu	49.18	—
Lian Zi	49.18	—
Ji Xue Teng	49.32	—
Ma Bian Cao	50.00	—

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
62604	62604	0.00	6.35
62352	62352	0.00	19.72
62603	62603	0.00	6.41
62726	62726	0.00	14.66
62351	62351	0.00	20.40
62725	62725	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 2.1-2020-01

Validated substance/substance group **Xu Duan**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60245-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Xu Duan	2	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	G092H2106821	62329	40	from supplier
PhytoComm	Xu Duan	G092H2106821	62330	40	from supplier
PhytoComm	Xu Duan	G092H2106822	62525	40	from supplier
PhytoComm	Xu Duan	G092H2106822	62526	40	from supplier

Validation samples

A total of 39 821 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 *Xu Duan*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G092H2106821	62329 [†]	20
PhytoComm	Xu Duan	G092H2106821	62330 [†]	20
PhytoComm	Xu Duan	G092H2106822	62525 [†]	20
PhytoComm	Xu Duan	G092H2106822	62526 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 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 *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	160	0	25 641
Type B	0	80	0	13 087
Type C	4	1	0	848

The substance/substance group *Xu Duan* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	98.9922 % (> 98.4094 %)	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 *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
Yin Chen Hao	11.69	—
Nü Zhen Zi	11.87	—
Wei Ling Xian	13.11	—
Ma Bian Cao	14.77	—
Huang Bo	16.36	—
Sang Bai Pi	16.41	—
Yi Mu Cao	16.64	—
Yin Yang Huo	17.27	—
Qiang Huo	17.45	—
Dan Zhu Ye	18.70	—
Ku Shen	21.54	—
Bu Gu Zhi	21.70	—
Gui Zhi	22.24	—
Dan Shen	22.54	—
Sang Ji Shend	22.99	—
Jing Jie	23.20	—
He Huan Pi	23.39	—
Gu Sui Bu	24.88	—
Wu Zhu Yu	25.51	—
Cang Er Zi	25.75	—
Yu Jin	25.84	—
Chen Pi	26.38	—
Fu Zi	27.18	—
He Shou Wu	27.22	—
Hou Po	27.41	—
Sha Ren	27.53	—
Lian Qiao	27.58	—
Hong Jing Tian	28.04	—
Xian He Cao	28.27	—
(Shi) Chang Pu	28.44	—
Fang Feng	28.90	—
Ze Xie	30.18	—
Qing Pi	31.10	—
Wu Yao	31.34	—
Gou Teng	31.79	—
Jiang Huang	31.82	—
Rou Cong Rong	32.69	—
Zi Wan	33.68	—
Du Zhong	33.98	—
Ju Hua	34.08	—
Gua Lou	34.15	—
Pu Gong Ying	36.33	—
She Gan	36.64	—
Gan Cao	36.99	—
Xie Bai	38.01	—
Ge Gen	38.50	—
Xiang Fu	38.72	—
Hu Zhang	39.38	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Lian	40.01	—
Mang Xiao	40.28	—
(Sheng) Di Huang	40.29	—
Huang Jing	40.81	—
Shen Qu	41.46	—
Chai Hu	41.56	—
Cang Zhu	41.87	—
Xin Yi	42.52	—
(Bai) Dou Kou	43.49	—
Wu Jia Pi	43.51	—
Ji Li	44.21	—
Xia Ku Cao	44.51	—
Zhi Zi	44.58	—
Niu Bang Zi	45.11	—
Lu Gen	45.48	—
Gou Qi Zi	45.88	—
Wu Wei Zi	46.27	—
Jin Qian Cao	46.55	—
Zhi Shi	46.61	—
Huang Qi	46.95	—
Sang Zhi	47.20	—
Huo Ma Ren	47.75	—
Chi Shao (Yao)	48.23	—
Du Huo	48.57	—
Han Lian Cao	48.61	—
Dang Gui	48.67	—
Shan Yao	49.31	—
Xi Xian Cao	49.44	—
Sang Ye	49.58	—

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
62526	62526	0.00	13.11
62525	62525	0.00	13.39
62330	62330	0.00	12.19
62329	62329	0.00	11.69

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 2.1-2020-01

Validated substance/substance group **Xuan Shen**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60095-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Xuan Shen	3	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	G223HS092QH1	62199	40	from supplier
PhytoComm	Xuan Shen	G223HS092QH1	62200	40	from supplier
PhytoComm	Xuan Shen	G223HS092RN1	62273	40	from supplier
PhytoComm	Xuan Shen	G223HS092RN1	62274	40	from supplier
PhytoComm	Xuan Shen	G223HS092SH1	62553	40	from supplier
PhytoComm	Xuan Shen	G223HS092SH1	62554	40	from supplier

Validation samples

A total of 39 821 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 *Xuan Shen*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G223HS092QH1	62199 [†]	20
PhytoComm	Xuan Shen	G223HS092QH1	62200 [†]	20
PhytoComm	Xuan Shen	G223HS092RN1	62273 [†]	20
PhytoComm	Xuan Shen	G223HS092RN1	62274 [†]	20
PhytoComm	Xuan Shen	G223HS092SH1	62553 [†]	20
PhytoComm	Xuan Shen	G223HS092SH1	62554 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	g22h0537322	1
Phytocomm	Xuan Shen	G223H0537421	1

- 850 spectra from 12 *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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	0	3	850

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8257 %)	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 *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
Sha Ren	16.57	—
Du Zhong	23.99	—
Jing Jie	24.00	—
Yin Chen Hao	24.27	—
Pu Gong Ying	25.22	—
Bu Gu Zhi	25.64	—
Zhi Shi	26.82	—
Xia Ku Cao	27.67	—
Guang Huo Xiang	29.49	—
Wu Jia Pi	29.89	—
Yin Yang Huo	30.69	—
(Sheng) Di Huang	30.92	—
Sang Bai Pi	32.40	—
Qing Hao	33.82	—
Yi Mu Cao	35.51	—
Jin Qian Cao	35.53	—
Sang Ji Shend	37.12	—
Gui Zhi	38.72	—
Dan Zhu Ye	38.82	—
Huang Bo	41.10	—
Xian He Cao	41.78	—
Shu Di (Huang)	41.98	—
Dan Shen	43.39	—
Nü Zhen Zi	43.67	—
Xu Duan	44.85	—
Hu Zhang	45.03	—
Ma Bian Cao	45.34	—
Wu Zhu Yu	45.36	—
Hua Shi	46.18	—
Qiang Huo	48.12	—
Yu Xing Cao	50.25	—

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
62274	62274	0.00	25.22
62200	62200	0.00	19.59
62199	62199	0.00	18.09
62273	62273	0.00	25.51
62554	62554	0.00	16.76
62553	62553	0.00	16.57

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 2.1-2020-01

Validated substance/substance group **Yan Hu Suo**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60439-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Yan Hu Suo	2	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

Validation samples

A total of 39 821 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 *Yan Hu Suo*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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

- 13 087 spectra from a total of 318 batches from further 172 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	G081H0805521	1
PhytoComm	Yan Hu Suo	G081H0805223	1
PhytoComm	Yan Hu Suo	G081H0805223	1
PhytoComm	Yan Hu Suo	G081H0805521	1

- 847 spectra from 12 *Apo-Ident* customers from a total of 512 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	0	6	847

The substance/substance group *Yan Hu Suo* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8235 %)	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 *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	6.12	—
Shan Yao	9.41	—
Pi Pa Ye	10.51	—
Jiang Huang	11.15	—
Fang Ji	11.95	—
Ren Dong Teng	12.29	—
Gui Zhi	12.67	—
Chai Hu	12.90	—
Bai Xian Pi	13.13	—
(Bai) Dou Kou	13.17	—
Sang Zhi	13.55	—
Yu Jin	13.78	—
Bai Bian Dou	14.05	—
Che Qian Zi	14.39	—
Jin Yin Hua	14.47	—
Mu Dan Pi	14.71	—
Fu Zi	15.13	—
Yin Yang Huo	15.30	—
Dan Shen	15.36	—
Dan Dou Chi	15.52	—
Xiang Fu	16.70	—
Qing Pi	16.74	—
Qiang Huo	16.98	—
Yuan Zhi	17.09	—
Shen Qu	17.11	—
Gan Jiang	17.41	—
Tu Fu Ling	17.94	—
Zi Hua Di Ding	17.96	—
Ji Xue Teng	18.21	—
Jiao Gu Lan	18.26	—
Ma Huang	18.31	—
Chuan Xiong	18.44	—
Mao Dong Qing	18.60	—
Ling Zhi	18.66	—
Hong Jing Tian	18.70	—
Tian Hua Fen	18.87	—
Bai Shao Yao	18.89	—
Ji Li	18.95	—
Ye Jiao Teng	19.09	—
Lian Qiao	19.15	—
Lian Zi	19.32	—
Sha Ren	19.41	—
He Huan Pi	19.58	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Wu Wei Zi	19.77	—
Wu Yao	20.69	—
Yi Mu Cao	20.71	—
Lu Gen	21.47	—
Du Zhong	21.65	—
Huang Lian	21.77	—
Gua Lou	22.02	—
She Gan	22.91	—
Huo Ma Ren	22.95	—
Di Gu Pi	23.08	—
Tao Ren	23.58	—
Bai Zi Ren	23.76	—
Zi Su Zi	24.25	—
Cang Er Zi	24.25	—
Ze Xie	25.06	—
Zhi Ke	25.23	—
Chuan Lian Zi	25.54	—
(Shi) Chang Pu	25.90	—
Shan Yu Rou	25.93	—
Tai Zi Shen	25.99	—
Dang Gui	26.09	—
Ba Ji Tian	26.11	—
Zi Wan	26.14	—
Cang Zhu	26.34	—
Huang Bai	26.82	—
Ma Huang Gen	27.10	—
Gou Teng	27.21	—
Sha Shen (Bei)	27.24	—
Xie Bai	27.32	—
Sheng Jiang	27.37	—
Mang Xiao	27.41	—
Zhe Bei Mu	27.78	—
Wang Bu Liu Xing	27.91	—
E Zhu	28.04	—
Rou Gui	28.12	—
Gou Qi Zi	28.13	—
Zhi Mu	28.63	—
Ban Xia (Jiang)	29.19	—
(Huai) Niu Xi	29.21	—
Chen Pi	29.37	—
Pu Gong Ying	29.57	—
Xian He Cao	30.10	—
Du Huo	30.20	—
Fu Xiao Mai	30.35	—
Fu Ling	30.55	—
Ma Bian Cao	30.60	—
Bu Gu Zhi	30.79	—
Niu Bang Zi	30.85	—
Suan Zao Ren	31.73	—
San Qi	31.78	—
Da Zao	32.60	—
Xin Yi	32.83	—
Dang Gui Wei	32.90	—
Yi Yi Ren	33.02	—
Sang Ye	33.42	—
Huang Qin	33.56	—
Bai Zhu	33.61	—
Yu Zhu	34.27	—
Jie Geng	34.65	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ci Wu Jia	34.77	—
Bai Mao Gen	35.97	—
Dang Shen	36.10	—
Hu Zhang	36.45	—
Huang Bo	36.45	—
Chuang Mu Xiang	37.02	—
Ku Shen	37.25	—
Gan Cao	37.53	—
Xi Xian Cao	37.83	—
Fang Feng	37.88	—
Nü Zhen Zi	38.03	—
Long Dan (Cao)	38.65	—
Fu Shen	39.15	—
Sang Ji Shend	40.05	—
Zhu Ru	40.56	—
Gu Sui Bu	41.58	—
Xia Ku Cao	41.65	—
San Leng	41.75	—
Ju Hua	42.00	—
Sang Bai Pi	42.10	—
Wei Ling Xian	42.35	—
Huang Qi	42.35	—
Lai Fu Zi	42.39	—
Xu Duan	42.59	—
Chi Shao (Yao)	43.51	—
Wu Zhu Yu	44.10	—
Long Yan Rou	44.58	—
Zhu Ling	45.18	—
Hong Hua	45.98	—
He Shou Wu	47.01	—
Yin Chen Hao	47.14	—
Wu Mei	47.37	—
Fo Shou	47.40	—
Zhi Zi	48.76	—
Dong Gua Zi	48.79	—
Bai He	49.34	—

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	11.15
62372	62372	0.00	11.28
62677	62677	0.00	6.54
62678	62678	0.00	6.12

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 2.1-2020-01

Validated substance/substance group **Ye Jiao Teng**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60068-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ye Jiao Teng	3	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	G288H0885621	62131	40	from supplier
PhytoComm	Ye Jiao Teng	G288H0885621	62132	40	from supplier
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 39 821 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 *Ye Jiao Teng*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 *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	G288H0885621	62131 [†]	20
PhytoComm	Ye Jiao Teng	G288H0885621	62132 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	G288H0885522	2
Phytocomm	Ye Jiao Teng	g228h0085223	1
Phytocomm	Ye Jiao Teng	G288H0885223	1
PhytoComm	Ye Jiao Teng	G288H0885522	2

- 847 spectra from 12 *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 *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	3	236	4	25 558
Type B	0	117	3	13 047
Type C	0	0	6	847

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	99.9891 % (> 99.9741 %)	98.3333 % (> 97.0833 %)
Type B	100.0000 % (> 99.9421 %)	97.5000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8235 %)	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 *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	4.91	—
Bai Xian Pi	6.85	—
Tu Fu Ling	7.67	—
Dan Dou Chi	7.84	—
Yin Yang Huo	9.26	—
Rou Gui	9.69	—
Ren Dong Teng	10.74	—
Ling Zhi	12.86	—
Chen Pi	13.61	—
Fo Shou	13.62	—
Dan Shen	14.12	—
Yuan Zhi	14.25	—
Tai Zi Shen	14.54	—
Gou Teng	14.55	—
Mao Dong Qing	14.84	—
Bai Shao Yao	14.94	—
Ma Huang Gen	15.22	—
Huo Ma Ren	15.27	—
Yu Jin	15.30	—
Lian Zi	15.40	—
Shen Qu	15.64	—
Ma Huang	15.85	—
He Huan Pi	15.87	—
Tao Ren	16.06	—
Bai Bian Dou	16.23	—
Yan Hu Suo	17.06	—
Hou Po	17.69	—
Sheng Jiang	18.89	—
Jin Yin Hua	19.57	—
Gui Zhi	20.20	—
Qiang Huo	20.35	—
Chuan Xiong	20.64	—
Ci Wu Jia	20.92	—
Pi Pa Ye	21.21	—
Chai Hu	22.19	—
Zhe Bei Mu	22.37	—
Shan Yao	22.38	—
Long Yan Rou	22.45	—
Fu Shen	22.47	—
Fang Ji	22.75	—
Da Zao	23.13	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Mu Dan Pi	23.30	—
Ban Xia (Jiang)	23.36	—
Che Qian Zi	23.64	—
Lian Qiao	23.67	—
Yi Yi Ren	23.96	—
Di Gu Pi	24.07	—
Gua Lou	24.31	—
Zi Hua Di Ding	24.38	—
Zhi Ke	24.43	—
Tian Men Dong	24.70	—
Ji Li	24.77	—
Hong Jing Tian	25.00	—
Huang Lian	25.03	—
Fu Ling	25.05	—
Jiao Gu Lan	25.20	—
Tian Hua Fen	25.61	—
Fu Zi	26.21	—
Yu Zhu	26.38	—
Du Zhong	27.08	—
Cang Zhu	29.57	—
Fu Xiao Mai	29.69	—
Bai Zi Ren	30.63	—
Qing Pi	30.97	—
Zhu Ru	31.86	—
Wu Wei Zi	32.29	—
Gou Qi Zi	32.74	—
Dang Gui Wei	32.87	—
Huang Qin	32.98	—
Dang Shen	32.99	—
Ban Zhi Lian	33.49	—
Zhi Mu	34.06	—
Mang Xiao	34.77	—
San Qi	34.94	—
Shan Yu Rou	35.21	—
Sang Zhi	36.53	—
Jie Geng	36.90	—
Chi Shao (Yao)	38.10	—
Dang Gui	38.42	—
Huang Bai	39.10	—
Pu Gong Ying	39.69	—
Bai Zhu	39.86	—
Gan Cao	40.17	—
Bai He	41.10	—
Ze Xie	42.66	—
Guang Huo Xiang	45.40	—
Hua Shi	45.71	—
San Leng	46.12	—
Suan Zao Ren	47.40	—
Ai Ye	47.48	—
Chuang Mu Xiang	47.59	—
Wang Bu Liu Xing	47.98	—
Chuan Lian Zi	49.59	—
Bai Mao Gen	50.59	—

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
62131	62131	0.00	40.87
62421	62421	0.00	4.91
62132	62132	0.00	40.71
62422	62422	0.00	5.28
62625	62625	0.00	7.14
62626	62626	0.00	6.85

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 2.1-2020-01

Validated substance/substance group **Yi Mu Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60105-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Yi Mu Cao	3	0	3

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	G134H1050621	61983	40	from supplier
PhytoComm	Yi Mu Cao	G134H1050621	61984	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G134H1050621	61983 [†]	20
PhytoComm	Yi Mu Cao	G134H1050621	61984 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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 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	Yi Mu Cao	g134h1050024	1
Phytocomm	Yi Mu Cao	G134H1050521	1
PhytoComm	Yi Mu Cao	G134H1050321	1
PhytoComm	Yi Mu Cao	G134H1050521	3
PhytoComm	Yi Mu Cao	G134H1050621	1

- 846 spectra from 12 *Apo-Ident* customers from a total of 512 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	25 561
Type B	0	120	0	13 047
Type C	2	1	6	844

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 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	99.8062 % (> 99.2178 %)	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
Gui Zhi	8.92	—
Xu Duan	13.14	—
Qiang Huo	14.82	—
Huang Bo	15.07	—
Wei Ling Xian	16.24	—
Yin Yang Huo	17.42	—
Nü Zhen Zi	17.73	—
Fu Zi	17.83	—
Wu Yao	17.84	—
Dan Zhu Ye	19.99	—
Bu Gu Zhi	20.56	—
Cang Er Zi	21.20	—
Qing Pi	21.62	—
Jiang Huang	21.81	—
Ze Xie	24.31	—
Sha Ren	24.72	—
He Huan Pi	24.79	—
Dan Shen	25.56	—
Ju Hua	26.53	—
Xin Yi	26.95	—
Hu Zhang	27.08	—
Gou Teng	27.28	—
Ma Bian Cao	27.46	—
Gua Lou	27.49	—
Rou Cong Rong	28.46	—
Pu Gong Ying	29.81	—
Xiang Fu	29.85	—
Hou Po	30.31	—
Shan Yao	30.45	—
Ku Shen	31.77	—
(Shi) Chang Pu	31.92	—
Hong Jing Tian	32.10	—
Yin Chen Hao	32.38	—
Shen Qu	33.03	—
Gu Sui Bu	33.16	—
Jing Jie	33.37	—
Huang Lian	33.54	—
Sang Ye	33.91	—
Sang Ji Shend	33.92	—
Lian Qiao	34.07	—
Niu Bang Zi	34.66	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chen Pi	35.11	—
Lu Gen	35.88	—
Chai Hu	36.22	—
Gou Qi Zi	36.51	—
Xi Xian Cao	36.64	—
Wu Wei Zi	37.00	—
Bai Hua She She Cao	37.03	—
Wu Zhu Yu	37.23	—
Han Lian Cao	37.34	—
(Bai) Dou Kou	37.61	—
She Gan	37.92	—
Sang Bai Pi	38.20	—
Fang Feng	38.97	—
Ji Li	38.99	—
Xie Bai	39.02	—
Mang Xiao	39.32	—
Zhi Zi	39.64	—
Du Zhong	39.87	—
Xia Ku Cao	40.66	—
Gan Cao	40.94	—
Qing Hao	41.17	—
Xian He Cao	41.81	—
Di Gu Pi	42.14	—
He Shou Wu	42.59	—
Cang Zhu	43.16	—
Dong Kui Zi	44.37	—
Mao Dong Qing	44.65	—
Yu Jin	44.75	—
Du Huo	45.49	—
Bai Mao Gen	45.94	—
Tu Fu Ling	46.45	—
Zi Wan	46.55	—
E Zhu	46.97	—
Zhi Shi	47.35	—
Gan Jiang	47.68	—
Jin Yin Hua	47.87	—
Yan Hu Suo	48.09	—
Wu Mei	48.75	—
Dang Gui	48.81	—
(Sheng) Di Huang	48.82	—
Zhi Mu	49.01	—
Sang Zhi	49.72	—

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
62542	62542	0.00	15.27
61984	61984	0.00	8.92
62410	62410	0.00	13.14
62409	62409	0.00	13.14
62541	62541	0.00	14.82
61983	61983	0.00	9.19

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 2.1-2020-01

Validated substance/substance group **Yi Yi Ren**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60174-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G078H1701622	62097	40	from supplier
PhytoComm	Yi Yi Ren	G078H1701622	62098	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G078H1701622	62097 [†]	20
PhytoComm	Yi Yi Ren	G078H1701622	62098 [†]	20
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

- 13 047 spectra from a total of 317 batches from further 172 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	G078H1701522	3
Phytocomm	Yi Yi Ren	g078h1701323	1
PhytoComm	Yi Yi Ren	G078H1701323	6

- 842 spectra from 12 *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 *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	13	233	7	25 548
Type B	17	116	4	13 030
Type C	0	8	3	842

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	99.9637 % (> 99.9487 %)	97.0833 % (> 95.8333 %)
Type B	99.9079 % (> 99.8790 %)	96.6667 % (> 94.1667 %)
Type C	100.0000 % (> 98.8225 %)	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	2.22	—
Zhu Ru	3.50	—
Sheng Jiang	5.57	—
Rou Gui	6.94	—
Fu Ling	7.66	—
Ban Xia (Jiang)	7.91	—
Di Gu Pi	8.23	—
Gui Zhi	8.28	—
Wang Bu Liu Xing	8.90	—
Bai Bian Dou	9.45	—
Ci Wu Jia	9.48	—
Fu Shen	9.71	—
He Huan Pi	10.23	—
Lian Zi	11.07	—
Ma Huang Gen	11.42	—
Ji Li	12.25	—
Tao Ren	12.28	—
Gou Teng	13.89	—
Bai Shao Yao	14.49	—
Ji Xue Teng	14.57	—
Shen Qu	14.81	—
Bai Zi Ren	15.04	—
Ling Zhi	15.19	—
Zhe Bei Mu	15.99	—
Huo Ma Ren	18.29	—
Bai Xian Pi	19.02	—
Tian Hua Fen	19.66	—
Tai Zi Shen	20.38	—
Gua Lou	21.25	—
Shan Yao	23.15	—
Yuan Zhi	23.35	—
Fo Shou	23.84	—
Yu Jin	23.85	—
Ye Jiao Teng	25.18	—
Pi Pa Ye	25.23	—
Tian Men Dong	25.39	—
Long Yan Rou	25.86	—
Da Zao	27.16	—
Chen Pi	27.18	—
Lian Qiao	27.25	—
Ren Dong Teng	27.64	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yan Hu Suo	29.23	—
Chuan Xiong	29.87	—
Tu Fu Ling	30.31	—
Dang Shen	33.21	—
Dan Shen	33.87	—
Shan Yu Rou	34.25	—
Fu Zi	34.52	—
Zhi Mu	34.56	—
Fang Ji	34.65	—
Mao Dong Qing	34.70	—
Wu Wei Zi	35.07	—
Dan Dou Chi	36.22	—
San Qi	37.09	—
Jiao Gu Lan	37.22	—
Dang Gui Wei	37.87	—
Mang Xiao	38.51	—
Yin Yang Huo	39.33	—
Chai Hu	39.60	—
Huang Qin	40.40	—
Jin Yin Hua	40.62	—
Zhi Ke	40.63	—
Che Qian Zi	42.32	—
Yu Zhu	43.74	—
Dang Gui	43.85	—
Hou Po	44.54	—
Cang Zhu	46.57	—
Suan Zao Ren	47.42	—
Ma Huang	47.55	—
Hong Jing Tian	48.76	—
Jie Geng	50.12	—

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
62097	62097	0.00	50.98
62098	62098	0.00	49.98
62280	62280	0.00	3.93
62279	62279	0.00	3.12
62676	62676	0.00	2.22
62675	62675	0.00	2.38

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 2.1-2020-01

Validated substance/substance group **Yin Chen Hao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60163-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G034H1030621	61832	40	from supplier
PhytoComm	Yin Chen Hao	G034H1030721	62397	40	from supplier
PhytoComm	Yin Chen Hao	G034H1030721	62398	40	from supplier

Validation samples

A total of 39 821 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 *Yin Chen Hao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 681 spectra from a total of 317 batches from further 172 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.

- 60 spectra of 3 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	G034H1030621	61832 [†]	20
PhytoComm	Yin Chen Hao	G034H1030721	62397 [†]	20
PhytoComm	Yin Chen Hao	G034H1030721	62398 [†]	20

- 13 107 spectra from a total of 318 batches from further 172 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

- 851 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 *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	4	120	0	25 677
Type B	8	60	0	13 099
Type C	3	0	2	848

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	99.9903 % (> 99.9754 %)	100.0000 % (> 95.0000 %)
Type B	99.9612 % (> 99.9324 %)	100.0000 % (> 90.0000 %)
Type C	99.7209 % (> 99.1349 %)	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 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
Du Zhong	7.65	—
Jing Jie	10.95	—
Bu Gu Zhi	11.22	—
Nü Zhen Zi	11.53	—
Dan Zhu Ye	12.88	—
Xu Duan	13.53	—
Ma Bian Cao	13.84	—
Sang Bai Pi	15.45	—
Qiang Huo	17.89	—
Gou Teng	19.58	—
Dan Shen	20.01	—
Yin Yang Huo	21.23	—
Sha Ren	21.78	—
Sang Ji Shend	22.30	—
Pu Gong Ying	23.03	—
Hou Po	23.32	—
Wu Zhu Yu	25.17	—
Hong Jing Tian	25.61	—
He Shou Wu	25.70	—
Yi Mu Cao	26.46	—
Xian He Cao	27.60	—
Wu Jia Pi	28.44	—
(Sheng) Di Huang	28.64	—
Gui Zhi	28.82	—
Huang Bo	29.12	—
Huang Lian	29.97	—
Gu Sui Bu	29.99	—
He Huan Pi	30.79	—
Wei Ling Xian	32.14	—
Yu Jin	32.47	—
Jin Qian Cao	32.53	—
Chen Pi	33.22	—
Fang Feng	33.94	—
Fu Zi	34.17	—
Xuan Shen	35.47	—
Zhi Shi	37.11	—
Wu Yao	40.11	—
Qing Hao	40.21	—
Yu Xing Cao	41.45	—
Ku Shen	41.66	—
Ge Gen	42.68	—
Xin Yi	42.80	—
Xia Ku Cao	44.34	—
Jiang Huang	44.84	—
Ze Xie	45.07	—
Ju Hua	45.15	—
Hu Zhang	45.47	—
Bai Hua She She Cao	46.58	—
Han Lian Cao	47.14	—
Mang Xiao	47.25	—
Hua Shi	48.99	—
Xiang Fu	49.51	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Cang Er Zi	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
62398	62398	0.00	12.03
61832	61832	0.00	7.65
62397	62397	0.00	11.53

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 2.1-2020-01

Validated substance/substance group **Yin Yang Huo**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60228-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Yin Yang Huo	2	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	G098H1114821	62327	40	from supplier
PhytoComm	Yin Yang Huo	G098H1114821	62328	40	from supplier
PhytoComm	Yin Yang Huo	G098HS241SH1	62683	40	from supplier
PhytoComm	Yin Yang Huo	G098HS241SH1	62684	40	from supplier

Validation samples

A total of 39 821 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 Yang Huo*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G098H1114821	62327 [†]	20
PhytoComm	Yin Yang Huo	G098H1114821	62328 [†]	20
PhytoComm	Yin Yang Huo	G098HS241SH1	62683 [†]	20
PhytoComm	Yin Yang Huo	G098HS241SH1	62684 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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

- 851 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 *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	0	160	0	25 641
Type B	0	78	2	13 087
Type C	0	0	2	851

The substance/substance group *Yin Yang 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	97.5000 % (> 93.7500 %)
Type C	100.0000 % (> 98.8279 %)	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 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	4.27	—
Mao Dong Qing	5.24	—
Hou Po	6.79	—
Ling Zhi	7.04	—
Yu Jin	7.54	—
Dan Dou Chi	7.86	—
Qiang Huo	9.73	—
Che Qian Zi	10.34	—
Ye Jiao Teng	11.00	—
Yan Hu Suo	11.36	—
Ren Dong Teng	11.61	—
Bai Bian Dou	11.74	—
Yuan Zhi	12.21	—
Tu Fu Ling	12.40	—
Shen Qu	12.44	—
Yi Mu Cao	12.71	—
Du Zhong	14.01	—
Lian Zi	14.09	—
Shan Yao	14.80	—
Fang Ji	15.17	—
Chai Hu	15.92	—
Zi Hua Di Ding	16.12	—
Pi Pa Ye	16.44	—
Mu Dan Pi	17.02	—
Dan Shen	17.46	—
Chuan Xiong	17.57	—
Ji Xue Teng	17.89	—
Huang Bo	18.03	—
Ma Huang	18.27	—
Ma Huang Gen	18.43	—
Rou Gui	18.66	—
Bu Gu Zhi	18.66	—
Qing Pi	18.94	—
Ji Li	19.01	—
Gui Zhi	20.80	—
Jin Yin Hua	21.14	—
Jiao Gu Lan	21.19	—
He Huan Pi	21.41	—
Hong Jing Tian	21.44	—
Gou Teng	21.62	—
Lian Qiao	21.96	—
Dan Zhu Ye	22.23	—
Ban Xia (Jiang)	22.77	—
Huang Lian	22.80	—
Bai Shao Yao	22.95	—
Chen Pi	23.58	—
Sheng Jiang	24.23	—
Nü Zhen Zi	24.44	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Xiao Mai	24.80	—
Huo Ma Ren	25.05	—
Gua Lou	25.87	—
Yi Yi Ren	25.88	—
Fu Zi	26.34	—
Yin Chen Hao	26.76	—
Zhi Ke	26.88	—
Tian Hua Fen	28.29	—
Wu Wei Zi	28.31	—
Di Gu Pi	28.33	—
Ma Bian Cao	28.55	—
Cang Zhu	28.88	—
Huang Bai	28.95	—
Fu Ling	29.01	—
Hu Zhang	29.09	—
Xia Ku Cao	29.17	—
Tao Ren	29.47	—
Xu Duan	30.57	—
Jiang Huang	30.84	—
Han Lian Cao	31.20	—
Zhe Bei Mu	31.39	—
Pu Gong Ying	31.48	—
Gou Qi Zi	31.55	—
Da Zao	31.90	—
Wei Ling Xian	32.96	—
Sang Ji Shend	32.97	—
Tai Zi Shen	33.71	—
Fu Shen	34.65	—
Sang Bai Pi	34.68	—
Mang Xiao	35.25	—
Ci Wu Jia	36.18	—
Ban Zhi Lian	36.61	—
Long Yan Rou	36.65	—
Bai Zi Ren	36.69	—
Sha Ren	36.80	—
Gu Sui Bu	37.54	—
Wu Zhu Yu	37.81	—
Xian He Cao	38.86	—
Zhu Ru	39.06	—
Sang Zhi	39.54	—
Fo Shou	39.56	—
Jing Jie	39.76	—
Bai Hua She She Cao	39.95	—
Xin Yi	40.03	—
Xi Xian Cao	40.53	—
He Shou Wu	40.60	—
Suan Zao Ren	40.74	—
Jie Geng	40.83	—
Chi Shao (Yao)	41.14	—
Wang Bu Liu Xing	41.23	—
San Qi	41.28	—
Zhi Shi	41.98	—
Wu Yao	42.04	—
Tian Men Dong	42.36	—
Ze Xie	43.19	—
Yu Zhu	43.24	—
Huang Qin	43.30	—
Shan Yu Rou	43.98	—
Chuang Mu Xiang	44.46	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Sang Ye	44.53	—
Bai He	45.12	—
Zhi Mu	45.58	—
Dang Gui Wei	45.72	—
Dang Gui	45.74	—
(Sheng) Di Huang	45.82	—
Cang Er Zi	46.25	—
Rou Cong Rong	46.32	—
E Zhu	48.19	—
Zi Su Zi	49.23	—
Dang Shen	49.59	—
Bai Zhu	49.99	—

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
62328	62328	0.00	12.88
62684	62684	0.00	4.27
62683	62683	0.00	4.34
62327	62327	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 2.1-2020-01

Validated substance/substance group **Yu Jin**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60171-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Yu Jin	2	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
PhytoComm	Yu Jin	G084H2201821	62517	40	from supplier
PhytoComm	Yu Jin	G084H2201821	62518	40	from supplier

Validation samples

A total of 39 821 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 Jin*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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
PhytoComm	Yu Jin	G084H2201821	62517 [†]	20
PhytoComm	Yu Jin	G084H2201821	62518 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	160	0	25 641
Type B	3	80	0	13 084
Type C	0	0	0	853

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 % (> 96.2500 %)
Type B	99.9830 % (> 99.9541 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8323 %)	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 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.

Substance	Distance in main model	Distance in second-stage model
Mao Dong Qing	4.74	—
Ling Zhi	5.53	—
Hou Po	6.52	—
Du Zhong	7.09	—
Yin Yang Huo	7.13	—
Qiang Huo	7.33	—
Bai Xian Pi	8.78	—
Dan Dou Chi	10.60	—
Xian He Cao	10.61	—
Fang Ji	10.62	—
Wu Yao	10.86	—
Che Qian Zi	11.48	—
Yan Hu Suo	12.89	—
Ren Dong Teng	13.58	—
Shen Qu	14.00	—
Shan Yao	14.89	—
Zi Hua Di Ding	14.89	—
Bai Bian Dou	14.94	—
Chai Hu	14.99	—
Qing Pi	15.08	—
Gui Zhi	15.37	—
He Huan Pi	15.40	—
Sha Ren	15.41	—
Ye Jiao Teng	15.94	—
Ji Li	16.05	—
Fu Zi	16.28	—
Chen Pi	17.26	—
Dan Shen	17.27	—
Mu Dan Pi	17.37	—
Pi Pa Ye	17.50	—
Jiao Gu Lan	17.91	—
Yuan Zhi	19.16	—
Lian Zi	19.22	—
Jiang Huang	19.26	—
Chuan Xiong	19.50	—
Yi Mu Cao	19.78	—
Hong Jing Tian	19.94	—
Ji Xue Teng	19.99	—
He Shou Wu	21.04	—
Lian Qiao	21.24	—
Gu Sui Bu	21.55	—
Lu Gen	21.83	—
Ma Huang	21.99	—
Ma Bian Cao	22.58	—
Huang Lian	22.73	—
Fang Feng	23.25	—
Xu Duan	23.29	—
(Bai) Dou Kou	23.49	—
Tu Fu Ling	23.87	—
Jin Yin Hua	24.05	—
Ban Xia (Jiang)	24.38	—
Sang Ji Shend	24.54	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang Gen	25.02	—
Huang Bai	25.29	—
Gou Teng	25.42	—
Mang Xiao	25.99	—
Sang Bai Pi	26.02	—
Sang Zhi	26.21	—
Wu Wei Zi	26.46	—
Bai Shao Yao	27.17	—
Gan Jiang	27.34	—
Zhi Ke	27.41	—
Sheng Jiang	28.09	—
Rou Gui	28.20	—
Gua Lou	28.21	—
Bu Gu Zhi	28.29	—
Yin Chen Hao	28.73	—
Pu Gong Ying	29.06	—
Xi Xian Cao	29.66	—
Di Gu Pi	30.27	—
Fu Xiao Mai	30.33	—
Huo Ma Ren	30.47	—
Huang Jing	30.68	—
Wu Zhu Yu	31.10	—
Cang Zhu	32.10	—
Yi Yi Ren	32.56	—
Zi Su Zi	32.66	—
Tian Hua Fen	32.75	—
Tao Ren	33.22	—
Nü Zhen Zi	33.54	—
Gou Qi Zi	33.94	—
Xiang Fu	33.98	—
Huang Bo	34.01	—
Dan Zhu Ye	34.98	—
Fu Ling	35.74	—
Wang Bu Liu Xing	35.87	—
Tai Zi Shen	35.88	—
(Shi) Chang Pu	35.96	—
Zi Wan	36.92	—
Ba Ji Tian	36.98	—
E Zhu	37.01	—
Sang Ye	37.36	—
Zhe Bei Mu	37.45	—
She Gan	37.53	—
Ge Gen	37.90	—
Suan Zao Ren	39.69	—
Da Zao	39.72	—
Hu Zhang	39.74	—
Xia Ku Cao	39.96	—
Du Huo	40.18	—
Bai Zi Ren	40.61	—
Dang Gui Wei	40.73	—
Ban Zhi Lian	40.74	—
Cang Er Zi	41.27	—
Shan Yu Rou	41.62	—
Ci Wu Jia	41.64	—
Fu Shen	41.94	—
Xie Bai	41.98	—
Dang Gui	42.09	—
San Qi	42.40	—
Ze Xie	42.76	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chuan Lian Zi	43.18	—
Bai Mao Gen	44.11	—
San Leng	44.30	—
Ju Hua	44.49	—
Wu Mei	44.79	—
Zhu Ru	44.95	—
Long Dan (Cao)	45.33	—
Chi Shao (Yao)	45.97	—
Xin Yi	46.38	—
Jie Geng	46.44	—
Ku Shen	46.45	—
Huang Qin	46.76	—
Huang Qi	47.64	—
Yu Zhu	47.81	—
Chuang Mu Xiang	48.94	—
(Huai) Niu Xi	49.15	—
Zhi Shi	49.21	—
Niu Bang Zi	49.88	—
Hong Hua	50.26	—

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
62517	62517	0.00	10.61
62278	62278	0.00	5.68
62277	62277	0.00	4.74
62518	62518	0.00	10.86

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 2.1-2020-01

Validated substance/substance group **Yu Xing Cao**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60606-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Yu Xing Cao; Houlttuyniae herba

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Yu Xing Cao	1	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

Validation samples

A total of 39 821 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 Xing Cao*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 848 spectra from 12 *Apo-Ident* customers from a total of 512 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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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 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
(Sheng) Di Huang	12.46	—
Wu Jia Pi	13.05	—
Sha Ren	28.26	—
Shu Di (Huang)	28.92	—
Pu Gong Ying	30.89	—
Ai Ye	31.39	—
Jing Jie	33.58	—
Yin Chen Hao	36.01	—
Sang Bai Pi	36.61	—
Dan Zhu Ye	38.05	—
Xuan Shen	39.22	—
Bu Gu Zhi	39.43	—
Guang Huo Xiang	43.01	—
Hua Shi	44.95	—
Du Zhong	47.58	—
Jue Ming Zi	47.70	—
Ge Gen	48.14	—

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
62584	62584	0.00	12.46
62583	62583	0.00	12.75

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 2.1-2020-01

Validated substance/substance group **Yu Zhu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60056-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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

- 13 127 spectra from a total of 319 batches from further 172 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	G195H0517421	2
PhytoComm	Yu Zhu	G195H0517121	1

- 850 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 *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	80	0	25 721
Type B	0	40	0	13 127
Type C	0	0	3	850

The substance/substance group *Yu Zhu* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8257 %)	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
Chen Pi	8.18	—
Gua Lou	8.42	—
Dang Gui Wei	8.70	—
Gou Qi Zi	11.91	—
Chuan Xiong	12.09	—
Dang Shen	12.53	—
Zhe Bei Mu	13.95	—
Tai Zi Shen	14.14	—
Zhi Mu	14.18	—
Cang Zhu	14.40	—
Yuan Zhi	14.60	—
Pi Pa Ye	14.78	—
Dang Gui	14.80	—
Long Yan Rou	15.71	—
San Qi	15.86	—
Bai Shao Yao	15.92	—
Ye Jiao Teng	16.34	—
Da Zao	17.54	—
Di Gu Pi	18.01	—
Tao Ren	18.72	—
He Huan Pi	18.98	—
Fo Shou	19.13	—
Ma Huang	19.37	—
Lian Qiao	19.51	—
Zhi Ke	19.79	—
Shen Qu	19.89	—
Lian Zi	20.09	—
Ren Dong Teng	20.13	—
Tian Men Dong	21.20	—
Hong Jing Tian	21.48	—
Ci Wu Jia	21.95	—
Chuang Mu Xiang	22.31	—
Rou Gui	22.61	—
Bai Bian Dou	22.69	—
Dan Shen	22.79	—
Bai Zi Ren	22.99	—
Ji Xue Teng	23.11	—
Gui Zhi	23.15	—
Ji Li	23.44	—
Shan Yao	23.76	—
Tu Fu Ling	23.91	—
Shan Yu Rou	23.95	—
Gou Teng	24.02	—
Bai Xian Pi	24.40	—
Sheng Jiang	24.49	—
Jiao Gu Lan	24.65	—
Jin Yin Hua	25.21	—
Hou Po	26.78	—
Jie Geng	26.86	—
Tian Hua Fen	26.87	—
Bai Zhu	27.11	—
Fu Zi	27.64	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huo Ma Ren	27.80	—
Yin Yang Huo	28.28	—
Huang Qin	28.39	—
Gan Cao	28.95	—
Ling Zhi	29.70	—
Ban Xia (Jiang)	29.97	—
Wu Wei Zi	30.11	—
Mu Dan Pi	30.28	—
Fu Ling	30.59	—
Suan Zao Ren	30.79	—
Mao Dong Qing	31.38	—
Fu Xiao Mai	31.87	—
Ze Xie	32.23	—
Dan Dou Chi	33.58	—
Chai Hu	34.08	—
Sang Zhi	34.22	—
Mai Men Dong	34.90	—
Yan Hu Suo	35.03	—
Fang Ji	35.62	—
Yi Yi Ren	35.79	—
Zhu Ru	36.46	—
Fu Shen	36.90	—
Che Qian Zi	37.13	—
Ma Huang Gen	38.51	—
Zi Hua Di Ding	38.61	—
Chi Shao (Yao)	39.03	—
Mang Xiao	39.62	—
San Leng	39.70	—
Yu Jin	40.97	—
Qing Pi	42.72	—
Huang Qi	43.16	—
Chuan Lian Zi	43.19	—
Qiang Huo	44.17	—
Huang Lian	46.37	—
Cang Er Zi	46.55	—

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
62606	62606	0.00	8.18
62605	62605	0.00	8.42

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 2.1-2020-01

Validated substance/substance group **Yuan Zhi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60100-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G194HS306RP1	62431	40	from supplier
PhytoComm	Yuan Zhi	G194HS306RP1	62432	40	from supplier
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

Validation samples

A total of 39 821 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.
- 25 561 spectra from a total of 316 batches from further 172 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	G194HS306RP1	62431 [†]	20
PhytoComm	Yuan Zhi	G194HS306RP1	62432 [†]	20
PhytoComm	Yuan Zhi	G194HS306SH1	62557 [†]	20
PhytoComm	Yuan Zhi	G194HS306SH1	62558 [†]	20
PhytoComm	Yuan Zhi	G194H1406921	62743 [†]	20
PhytoComm	Yuan Zhi	G194H1406921	62744 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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	G194H1406522	2
PhytoComm	Yuan Zhi	G194H1406222	1
PhytoComm	Yuan Zhi	G194H1406522	1

- 847 spectra from 12 *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 *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	0	240	0	25 561
Type B	3	117	3	13 044
Type C	1	2	4	846

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	100.0000 % (> 99.9701 %)	100.0000 % (> 97.5000 %)
Type B	99.9855 % (> 99.9565 %)	97.5000 % (> 95.0000 %)
Type C	99.9225 % (> 99.3342 %)	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
Cang Zhu	4.44	—
Bai Shao Yao	5.16	—
Dang Shen	6.90	—
Lian Qiao	7.59	—
Suan Zao Ren	8.16	—
(Huai) Niu Xi	8.38	—
Chuan Xiong	8.79	—
Shan Yao	11.00	—
Jiao Gu Lan	11.09	—
Ji Li	11.46	—
Bai Bian Dou	12.04	—
Mao Dong Qing	12.11	—
Ye Jiao Teng	12.20	—
Pi Pa Ye	12.33	—
Chen Pi	12.44	—
Gua Lou	12.85	—
Gou Qi Zi	12.96	—
Zhi Mu	13.12	—
San Qi	13.17	—
Bai Zi Ren	13.24	—
Huang Qi	13.46	—
Di Gu Pi	13.49	—
Jin Yin Hua	13.55	—
Jie Geng	13.77	—
Fo Shou	13.94	—
Tian Men Dong	14.58	—
(Shi) Chang Pu	15.40	—
Shen Qu	15.58	—
Dan Shen	16.03	—
Mu Gua	16.11	—
Shan Yu Rou	16.28	—
He Huan Pi	16.31	—
Chuang Mu Xiang	16.37	—
Dang Gui Wei	16.55	—
Lian Zi	16.70	—
Ban Xia (Jiang)	16.76	—
Mu Dan Pi	17.81	—
Ling Zhi	18.05	—
Rou Gui	18.27	—
Ze Xie	18.52	—
Huo Ma Ren	18.61	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Fu Zi	18.87	—
Bai Mao Gen	19.15	—
Zi Wan	19.41	—
Wu Wei Zi	19.43	—
Chi Shao (Yao)	19.96	—
Sang Zhi	20.10	—
Hou Po	20.12	—
Gui Zhi	20.16	—
Dang Gui	20.24	—
Ci Wu Jia	20.36	—
Zhe Bei Mu	20.47	—
Tu Fu Ling	20.49	—
Yu Zhu	20.62	—
Fu Xiao Mai	20.64	—
She Gan	20.66	—
Ren Dong Teng	20.85	—
Zhu Ru	21.54	—
Qing Pi	21.77	—
Tai Zi Shen	22.22	—
Chai Hu	22.24	—
Fu Ling	22.31	—
Da Zao	22.56	—
San Leng	23.05	—
Tian Hua Fen	23.24	—
Hong Jing Tian	23.58	—
Sheng Jiang	23.60	—
Niu Bang Zi	23.76	—
Mai Men Dong	24.33	—
Tao Ren	25.12	—
Bai Xian Pi	25.94	—
Xiang Fu	26.18	—
Fang Ji	26.22	—
Ku Shen	26.42	—
Bai Zhu	26.66	—
Yan Hu Suo	27.46	—
Ji Xue Teng	27.69	—
Cang Er Zi	27.98	—
Dan Dou Chi	28.15	—
Gou Teng	28.21	—
Yin Yang Huo	28.55	—
Gan Cao	28.83	—
Zhi Ke	29.32	—
Yi Yi Ren	30.18	—
Huang Qin	31.35	—
Wei Ling Xian	31.49	—
Wang Bu Liu Xing	31.57	—
Qiang Huo	31.94	—
Yu Jin	32.35	—
Long Yan Rou	32.77	—
Zi Hua Di Ding	33.22	—
Du Zhong	33.35	—
Gan Jiang	33.35	—
Che Qian Zi	33.83	—
Mang Xiao	34.26	—
Yi Mu Cao	34.50	—
Du Huo	35.16	—
Ma Huang	35.20	—
Lai Fu Zi	35.22	—
Qin Jiao	36.46	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ma Huang Gen	39.56	—
Huang Lian	39.65	—
Chuan Lian Zi	40.47	—
Zhi Zi	40.87	—
Lu Gen	41.17	—
Fu Shen	41.96	—
Dong Gua Zi	42.18	—
Fang Feng	42.23	—
Ju Hua	42.50	—
Hong Hua	42.73	—
Sang Ye	43.59	—
Zi Su Zi	44.03	—
Wu Yao	45.04	—
Bai He	45.09	—
Sha Shen (Bei)	45.29	—
E Zhu	45.94	—
Jiang Huang	46.38	—
Huang Bai	47.30	—
Xie Bai	47.58	—
Long Dan (Cao)	48.73	—
Pu Gong Ying	48.87	—

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
62744	62744	0.00	8.58
62431	62431	0.00	4.44
62557	62557	0.00	5.16
62743	62743	0.00	8.38
62558	62558	0.00	5.56
62432	62432	0.00	5.08

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 2.1-2020-01

Validated substance/substance group **Ze Xie**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60165-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Ze Xie	2	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	G011H1605821	62295	40	from supplier
PhytoComm	Ze Xie	G011H1605821	62296	40	from supplier
PhytoComm	Ze Xie	G011H1605822	62507	40	from supplier
PhytoComm	Ze Xie	G011H1605822	62508	40	from supplier

Validation samples

A total of 39 821 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 *Ze Xie*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 *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	G011H1605821	62295 [†]	20
PhytoComm	Ze Xie	G011H1605821	62296 [†]	20
PhytoComm	Ze Xie	G011H1605822	62507 [†]	20
PhytoComm	Ze Xie	G011H1605822	62508 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G011H1605521	5
Phytocomm	Ze Xie	G011H1605422	1
PhytoComm	Ze Xie	G011H1605122	1
PhytoComm	Ze Xie	G011H1605521	1

- 844 spectra from 12 *Apo-Ident* customers from a total of 512 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	160	0	25 641
Type B	0	80	0	13 087
Type C	3	3	6	841

The substance/substance group *Ze Xie* 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.3023 % (> 98.7137 %)	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 *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
(Shi) Chang Pu	12.69	—
Xiang Fu	14.54	—
Qing Pi	14.83	—
Zi Wan	15.82	—
Dang Gui Wei	16.47	—
Sang Zhi	16.76	—
Bai Mao Gen	17.35	—
Gua Lou	17.58	—
Gan Cao	17.67	—
Lian Qiao	17.74	—
Chi Shao (Yao)	17.85	—
Yuan Zhi	18.00	—
Mu Dan Pi	19.05	—
Di Gu Pi	19.78	—
Yi Mu Cao	21.30	—
She Gan	21.33	—
Zhi Ke	22.57	—
San Leng	22.84	—
Gui Zhi	23.41	—
Niu Bang Zi	24.68	—
Shan Yao	25.32	—
Wu Wei Zi	25.88	—
Cang Er Zi	26.00	—
Wu Yao	27.54	—
Jie Geng	28.90	—
Huang Qi	28.95	—
(Huai) Niu Xi	29.27	—
Bai Shao Yao	29.38	—
Du Huo	29.48	—
Jin Yin Hua	29.58	—
Chuan Xiong	29.79	—
Gan Jiang	29.96	—
Zi Hua Di Ding	30.32	—
Ku Shen	30.64	—
Wei Ling Xian	30.66	—
Cang Zhu	30.98	—
Gou Qi Zi	31.41	—
Shan Yu Rou	31.73	—
Huang Lian	32.85	—
Pu Gong Ying	33.01	—
Huo Ma Ren	33.71	—
Qiang Huo	34.07	—
Hong Jing Tian	34.37	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Tian Hua Fen	34.70	—
Ma Huang	34.75	—
Sha Shen (Bei)	34.92	—
Chai Hu	35.00	—
Mang Xiao	35.07	—
Fu Zi	35.08	—
Chen Pi	35.26	—
Jiang Huang	35.64	—
Lu Gen	35.83	—
Bai Zhu	37.23	—
Huang Bai	37.53	—
Hou Po	38.03	—
Dan Shen	38.22	—
He Huan Pi	38.76	—
Jiao Gu Lan	38.84	—
Dang Gui	38.93	—
Zhi Zi	39.17	—
Ye Jiao Teng	39.19	—
Mao Dong Qing	39.27	—
Huang Qin	39.39	—
Yin Yang Huo	39.51	—
Dang Shen	39.77	—
Ban Zhi Lian	39.80	—
E Zhu	40.16	—
Ren Dong Teng	40.17	—
Mu Gua	40.54	—
Gou Teng	40.83	—
Dan Dou Chi	40.92	—
Lian Zi	41.36	—
Fang Ji	41.62	—
(Bai) Dou Kou	41.64	—
Sang Ye	42.02	—
Zhe Bei Mu	42.05	—
Tu Fu Ling	42.52	—
Qin Jiao	42.64	—
Xu Duan	43.44	—
Fang Feng	43.56	—
Chuan Lian Zi	43.58	—
Xin Yi	43.58	—
Lai Fu Zi	44.57	—
Huang Bo	44.63	—
Long Dan (Cao)	45.06	—
Bai Xian Pi	45.24	—
Che Qian Zi	45.52	—
Hong Hua	45.57	—
Dong Gua Zi	45.59	—
Du Zhong	45.70	—
Ji Li	45.74	—
Sha Ren	45.91	—
Sheng Jiang	45.94	—
Zhi Mu	46.66	—
Yan Hu Suo	46.70	—
Ji Xue Teng	46.90	—
Zi Su Zi	47.82	—
Ju Hua	48.13	—
Xi Xian Cao	48.49	—
Pi Pa Ye	49.10	—
San Qi	49.23	—

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
62508	62508	0.00	14.54
62296	62296	0.00	12.85
62295	62295	0.00	12.69
62507	62507	0.00	14.12

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 2.1-2020-01

Validated substance/substance group **Zhe Bei Mu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50388-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G108H0712821	62339	40	from supplier
PhytoComm	Zhe Bei Mu	G108H0712821	62340	40	from supplier
PhytoComm	Zhe Bei Mu	G108HS138SG1	62687	40	from supplier
PhytoComm	Zhe Bei Mu	G108HS138SG1	62688	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G108H0712821	62339 [†]	20
PhytoComm	Zhe Bei Mu	G108H0712821	62340 [†]	20
PhytoComm	Zhe Bei Mu	G108HS138SG1	62687 [†]	20
PhytoComm	Zhe Bei Mu	G108HS138SG1	62688 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	G108H0712422	1
PhytoComm	Zhe Bei Mu	G108H0712321	1

- 851 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 *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	25 641
Type B	0	80	0	13 087
Type C	0	0	2	851

The substance/substance group *Zhe Bei 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 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Tian Hua Fen	8.81	—
Pi Pa Ye	8.98	—
Tai Zi Shen	9.27	—
Bai Shao Yao	9.27	—
Lian Zi	9.64	—
Bai Bian Dou	10.28	—
He Huan Pi	10.40	—
Di Gu Pi	11.00	—
Bai Zi Ren	11.79	—
Shan Yao	11.82	—
Ban Xia (Jiang)	11.86	—
Lian Qiao	12.16	—
Shen Qu	13.05	—
Ye Jiao Teng	13.47	—
Gua Lou	13.50	—
Rou Gui	13.93	—
Tao Ren	14.21	—
Ji Li	14.53	—
Sheng Jiang	15.24	—
Ji Xue Teng	15.55	—
Chuan Xiong	15.78	—
Huo Ma Ren	17.00	—
Yi Yi Ren	17.10	—
Ci Wu Jia	17.12	—
Gui Zhi	17.33	—
Fu Zi	17.35	—
Zhu Ru	17.98	—
Jin Yin Hua	18.03	—
Yan Hu Suo	18.04	—
Yuan Zhi	18.18	—
Fu Xiao Mai	19.58	—
Gou Teng	21.19	—
Fu Ling	21.20	—
Dang Shen	21.65	—
Chen Pi	21.93	—
Ren Dong Teng	22.12	—
Ling Zhi	22.29	—
Shan Yu Rou	22.36	—
Fo Shou	22.75	—
Bai Xian Pi	22.99	—
Zhi Ke	23.26	—
Wu Wei Zi	23.69	—
Fang Ji	24.06	—
Jiao Gu Lan	24.11	—
Hou Po	24.17	—
Da Zao	24.30	—
Dang Gui Wei	24.42	—
Dang Gui	24.94	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Qin	25.08	—
Cang Zhu	25.46	—
Yu Zhu	26.36	—
San Qi	26.52	—
Ma Huang	27.35	—
Yu Jin	28.58	—
Zhi Mu	28.98	—
Long Yan Rou	29.60	—
Che Qian Zi	29.61	—
Yin Yang Huo	29.97	—
Tian Men Dong	31.45	—
Fu Shen	31.60	—
Ma Huang Gen	31.75	—
Dan Shen	32.09	—
Tu Fu Ling	32.95	—
Dan Dou Chi	33.94	—
Mao Dong Qing	35.15	—
Gou Qi Zi	36.02	—
Mu Dan Pi	36.04	—
Wang Bu Liu Xing	36.42	—
Wu Mei	36.94	—
Chai Hu	37.67	—
Gan Cao	38.02	—
Cang Er Zi	38.40	—
Huang Lian	40.40	—
Mang Xiao	40.89	—
Qing Pi	40.91	—
Bai Zhu	41.16	—
Hong Jing Tian	41.69	—
Sang Zhi	41.84	—
Zhi Zi	42.15	—
Suan Zao Ren	42.16	—
Qiang Huo	42.18	—
Yi Mu Cao	42.81	—
Ze Xie	44.49	—
Chi Shao (Yao)	44.71	—
Nü Zhen Zi	44.77	—
Chuang Mu Xiang	45.98	—
Jie Geng	46.88	—
Du Zhong	47.27	—
Dan Zhu Ye	50.58	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/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
62340	62340	0.00	37.09
62688	62688	0.00	9.27
62687	62687	0.00	8.81
62339	62339	0.00	36.94

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 2.1-2020-01

Validated substance/substance group **Zhi Ke**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60158-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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

- 13 087 spectra from a total of 318 batches from further 172 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

- 850 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 *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	25 641
Type B	0	80	0	13 087
Type C	1	0	3	849

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	99.7674 % (> 99.1803 %)	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	9.75	—
Ma Huang	12.19	—
Hou Po	13.39	—
Lian Qiao	13.58	—
Jiao Gu Lan	14.16	—
Wu Wei Zi	16.21	—
Zi Hua Di Ding	16.69	—
Shan Yu Rou	17.71	—
Bai Zhu	18.24	—
Hong Jing Tian	18.48	—
Pi Pa Ye	18.50	—
Gou Qi Zi	18.99	—
Fu Zi	19.00	—
Yan Hu Suo	19.14	—
San Qi	19.16	—
Fang Ji	19.29	—
Ye Jiao Teng	19.43	—
Yin Yang Huo	19.69	—
Dan Shen	19.98	—
Cang Zhu	20.29	—
Yu Zhu	20.41	—
Zhe Bei Mu	20.61	—
Gua Lou	21.03	—
Chuang Mu Xiang	21.15	—
Shan Yao	21.31	—
Bai Shao Yao	21.92	—
Qing Pi	22.10	—
Tai Zi Shen	22.68	—
Huang Qin	22.88	—
Dang Gui	23.28	—
Chai Hu	23.99	—
Gui Zhi	24.08	—
Mu Dan Pi	24.25	—
He Huan Pi	24.43	—
Dang Shen	24.48	—
Chuan Xiong	24.53	—
Gan Cao	25.57	—
Huang Bai	25.69	—
Qiang Huo	25.98	—
Zhi Mu	26.53	—
Ren Dong Teng	26.62	—
Huang Lian	26.74	—
Bai Bian Dou	27.26	—
Bai Xian Pi	28.44	—
Bai Zi Ren	28.58	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Yuan Zhi	28.71	—
Suan Zao Ren	28.85	—
Chen Pi	29.13	—
Di Gu Pi	29.40	—
Ji Li	29.94	—
Tao Ren	30.04	—
Sang Ye	30.36	—
Tian Hua Fen	30.65	—
Ji Xue Teng	30.91	—
Dang Gui Wei	31.80	—
Lian Zi	33.63	—
Du Zhong	34.12	—
Chi Shao (Yao)	34.94	—
Hong Hua	35.16	—
Cang Er Zi	35.61	—
Mao Dong Qing	35.72	—
Dan Dou Chi	36.16	—
Che Qian Zi	36.39	—
Sheng Jiang	37.45	—
Huo Ma Ren	37.96	—
Ci Wu Jia	39.64	—
Yu Jin	39.98	—
Tu Fu Ling	41.78	—
Ling Zhi	41.80	—
Yi Mu Cao	42.18	—
Bai Mao Gen	42.34	—
Rou Gui	42.48	—
Mang Xiao	42.66	—
Gou Teng	43.08	—
Ze Xie	43.33	—
Shen Qu	43.36	—
Fu Xiao Mai	43.89	—
Ban Xia (Jiang)	44.12	—
Yi Yi Ren	45.30	—
Pu Gong Ying	45.87	—
San Leng	46.78	—
Fu Ling	46.86	—
Wu Yao	46.89	—
Xia Ku Cao	47.08	—
Da Zao	47.38	—
Long Yan Rou	48.43	—
Ban Zhi Lian	48.76	—
Sang Zhi	49.20	—
Zi Wan	49.88	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/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
62704	62704	0.00	13.68
62440	62440	0.00	9.75
62703	62703	0.00	13.58
62439	62439	0.00	10.82

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 2.1-2020-01

Validated substance/substance group **Zhi Mu**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60196-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Zhi Mu	3	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	G019HS153QL1	62193	40	from supplier
PhytoComm	Zhi Mu	G019HS153QL1	62194	40	from supplier
PhytoComm	Zhi Mu	G019H0824821	62405	40	from supplier
PhytoComm	Zhi Mu	G019H0824821	62406	40	from supplier
PhytoComm	Zhi Mu	G019HS153SG1	62647	40	from supplier
PhytoComm	Zhi Mu	G019HS153SG1	62648	40	from supplier

Validation samples

A total of 39 821 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 Mu*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 3 different batches.
- 25 561 spectra from a total of 316 batches from further 172 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 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	G019HS153QL1	62193 [†]	20
PhytoComm	Zhi Mu	G019HS153QL1	62194 [†]	20
PhytoComm	Zhi Mu	G019H0824821	62405 [†]	20
PhytoComm	Zhi Mu	G019H0824821	62406 [†]	20
PhytoComm	Zhi Mu	G019HS153SG1	62647 [†]	20
PhytoComm	Zhi Mu	G019HS153SG1	62648 [†]	20

- 13 047 spectra from a total of 317 batches from further 172 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

- 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 *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	240	0	25 561
Type B	0	120	0	13 047
Type C	0	0	1	852

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 %)	100.0000 % (> 97.5000 %)
Type B	100.0000 % (> 99.9421 %)	100.0000 % (> 95.0000 %)
Type C	100.0000 % (> 98.8343 %)	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 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	9.31	—
Dang Shen	9.47	—
Gua Lou	10.44	—
Chuan Xiong	10.56	—
Bai Shao Yao	10.70	—
(Huai) Niu Xi	10.98	—
Gou Qi Zi	12.45	—
Chai Hu	12.47	—
Shan Yu Rou	12.48	—
Cang Zhu	13.92	—
She Gan	14.41	—
Shan Yao	14.45	—
Sang Zhi	14.55	—
Gan Jiang	15.25	—
Pi Pa Ye	15.71	—
Dang Gui Wei	15.82	—
San Qi	15.84	—
Mai Men Dong	16.96	—
Jie Geng	16.97	—
Chen Pi	17.02	—
Ye Jiao Teng	17.12	—
Sheng Jiang	17.90	—
Di Gu Pi	18.37	—
Bai Bian Dou	18.42	—
Ba Ji Tian	19.08	—
Tian Men Dong	19.63	—
Xie Bai	19.78	—
Mu Gua	19.85	—
(Shi) Chang Pu	20.42	—
Huang Qi	21.10	—
Bai Mao Gen	21.22	—
Lai Fu Zi	21.26	—
Long Dan (Cao)	21.36	—
Ji Li	21.42	—
Dang Gui	22.01	—
Lian Qiao	22.04	—
Qing Pi	22.23	—
Shen Qu	22.28	—
Zi Wan	22.76	—
Fo Shou	23.04	—
Bai Zhu	23.16	—
Lian Zi	23.17	—
He Huan Pi	23.52	—
Fang Feng	23.78	—
Ban Xia (Jiang)	24.09	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Zhi Ke	24.38	—
Bai Zi Ren	24.81	—
Du Huo	25.22	—
Tian Hua Fen	25.31	—
San Leng	25.74	—
Zi Su Zi	25.79	—
Gui Zhi	25.86	—
Rou Gui	26.01	—
Jiao Gu Lan	26.12	—
Suan Zao Ren	26.27	—
Mang Xiao	26.28	—
Bai He	26.29	—
Fu Xiao Mai	26.71	—
Niu Bang Zi	26.77	—
Zhe Bei Mu	26.82	—
Huo Ma Ren	27.22	—
Yu Zhu	27.22	—
Fu Ling	27.29	—
Tu Fu Ling	27.35	—
Tai Zi Shen	27.39	—
Ren Dong Teng	27.41	—
Zhu Ru	27.48	—
E Zhu	27.68	—
Mao Dong Qing	28.18	—
Da Zao	28.64	—
Mu Dan Pi	28.71	—
Gan Cao	29.26	—
Lu Gen	29.32	—
Jin Yin Hua	29.80	—
Ci Wu Jia	29.86	—
Chuang Mu Xiang	30.25	—
Yan Hu Suo	30.40	—
Hou Po	30.42	—
Wu Wei Zi	30.58	—
Chuan Lian Zi	30.90	—
Jiang Huang	31.11	—
Xiang Fu	31.11	—
Ling Zhi	31.22	—
Qin Jiao	31.27	—
Sha Shen (Bei)	32.02	—
Huang Qin	32.98	—
Ku Shen	33.01	—
Yi Mu Cao	33.57	—
Wang Bu Liu Xing	33.66	—
Cang Er Zi	33.66	—
Ji Xue Teng	33.88	—
Gou Teng	34.35	—
Tao Ren	34.63	—
Hong Jing Tian	34.75	—
Long Yan Rou	35.34	—
Ju Hua	35.78	—
Yi Yi Ren	36.53	—
Dan Shen	36.76	—
Chi Shao (Yao)	37.01	—
Fang Ji	37.88	—
Dong Gua Zi	38.44	—
Xian He Cao	38.63	—
Bai Xian Pi	38.99	—
Ma Huang	38.99	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Ze Xie	39.69	—
Fu Zi	39.77	—
Wu Yao	40.41	—
Sha Ren	40.81	—
Mai Ya	42.01	—
(Bai) Dou Kou	42.81	—
Yin Yang Huo	42.84	—
Che Qian Zi	44.21	—
Wei Ling Xian	44.54	—
Mo Yao	47.23	—
Zhi Zi	48.05	—

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
62405	62405	0.00	10.44
62648	62648	0.00	9.60
62194	62194	0.00	13.74
62193	62193	0.00	14.62
62647	62647	0.00	9.31
62406	62406	0.00	10.97

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 2.1-2020-01

Validated substance/substance group **Zhi Shi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60145-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Zhi Shi	2	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	G044H0922821	62389	40	from supplier
PhytoComm	Zhi Shi	G044H0922821	62390	40	from supplier
PhytoComm	Zhi Shi	G044HS168SK1	62655	40	from supplier
PhytoComm	Zhi Shi	G044HS168SK1	62656	40	from supplier

Validation samples

A total of 39 821 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 Shi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	G044H0922821	62389 [†]	20
PhytoComm	Zhi Shi	G044H0922821	62390 [†]	20
PhytoComm	Zhi Shi	G044HS168SK1	62655 [†]	20
PhytoComm	Zhi Shi	G044HS168SK1	62656 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	g044h0922321	1
Phytocomm	Zhi Shi	G044H0922221	2
Phytocomm	Zhi Shi	g044h0922221	1
Phytocomm	Zhi Shi	G044H0922321	2
Phytocomm	Zhi Shi	G044H0922521	3
PhytoComm	Zhi Shi	G044H0922321	2
PhytoComm	Zhi Shi	G044H0922221	3

- 839 spectra from 12 *Apo-Ident* customers from a total of 511 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	160	0	25 641
Type B	0	80	0	13 087
Type C	0	1	13	839

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 % (> 99.9701 %)	100.0000 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8223 %)	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
Xia Ku Cao	11.51	—
Hu Zhang	14.42	—
Bai Hua She She Cao	19.26	—
Ge Gen	20.32	—
Gui Zhi	25.54	—
Han Lian Cao	25.71	—
Xuan Shen	26.81	—
Sang Ye	27.03	—
Xin Yi	28.21	—
Yi Mu Cao	30.53	—
Huang Bo	30.66	—
Yin Yang Huo	30.94	—
Yin Chen Hao	31.08	—
Bu Gu Zhi	33.61	—
Qiang Huo	34.42	—
Sha Ren	36.46	—
Gu Sui Bu	36.77	—
Wu Zhu Yu	37.20	—
Huang Lian	37.24	—
Dan Zhu Ye	37.85	—
Dan Shen	38.27	—
Xi Xian Cao	38.38	—
Sang Ji Shend	39.30	—
Rou Cong Rong	40.31	—
He Huan Pi	40.65	—
He Shou Wu	41.01	—
Gou Teng	41.08	—
Xian He Cao	41.28	—
Jing Jie	41.40	—
Jiang Huang	41.54	—
Ma Huang	41.59	—
Xu Duan	42.41	—
Sang Bai Pi	43.70	—
Fu Zi	44.11	—
Hong Jing Tian	44.62	—
Pu Gong Ying	44.90	—
Wei Ling Xian	45.27	—
Shen Qu	47.02	—
Ma Bian Cao	47.59	—
Nü Zhen Zi	48.00	—
Hou Po	48.62	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Shan Yao	48.73	—

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
62656	62656	0.00	11.51
62390	62390	0.00	20.54
62389	62389	0.00	20.32
62655	62655	0.00	12.33

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 2.1-2020-01

Validated substance/substance group **Zhi Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60367-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Zhi Zi; Gardeniae fructus

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Zhi Zi	2	0	2

Second-stage model

For differentiation of the substance/substance group *Zhi 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 *Zhi Zi*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zhi Zi	G110H1137822	62529	40	from supplier
PhytoComm	Zhi Zi	G110H1137822	62530	40	from supplier
PhytoComm	Zhi Zi	G110H1137922	62733	40	from supplier
PhytoComm	Zhi Zi	G110H1137922	62734	40	from supplier

Validation samples

A total of 39 821 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 Zi*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 2 different batches.
- 25 641 spectra from a total of 317 batches from further 172 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 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	Zhi Zi	G110H1137822	62529 [†]	20
PhytoComm	Zhi Zi	G110H1137822	62530 [†]	20
PhytoComm	Zhi Zi	G110H1137922	62733 [†]	20
PhytoComm	Zhi Zi	G110H1137922	62734 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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 *Zhi 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	Zhi Zi	g110h1137121	1
Phytocomm	Zhi Zi	G110H1137521	3
PhytoComm	Zhi Zi	G110H1137121	1
PhytoComm	Zhi Zi	G110H1137521	1

- 847 spectra from 12 *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 *Zhi 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 *Zhi 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	25 641
Type B	0	80	0	13 087
Type C	0	0	6	847

The substance/substance group *Zhi 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 % (> 96.2500 %)
Type B	100.0000 % (> 99.9422 %)	100.0000 % (> 92.5000 %)
Type C	100.0000 % (> 98.8235 %)	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 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
Cang Er Zi	16.57	—
Ku Shen	17.66	—
Wei Ling Xian	23.46	—
Dong Gua Zi	24.84	—
Gua Lou	25.08	—
Huo Ma Ren	25.71	—
Niu Bang Zi	26.34	—
Shen Qu	27.87	—
Gui Zhi	28.99	—
Shan Yao	30.72	—
Lian Qiao	31.72	—
Di Gu Pi	35.67	—
Huang Bo	37.28	—
Qing Pi	38.11	—
She Gan	38.15	—
Jiang Huang	38.53	—
Nü Zhen Zi	38.68	—
Wu Wei Zi	39.00	—
Yi Mu Cao	39.04	—
Ze Xie	39.07	—
Suan Zao Ren	39.87	—
Xiang Fu	40.34	—
Jin Yin Hua	40.34	—
Sang Zhi	40.95	—
Mang Xiao	41.20	—
Bai Bian Dou	41.81	—
Gan Cao	42.96	—
Zi Wan	43.30	—
(Shi) Chang Pu	43.44	—
Tian Hua Fen	43.87	—
Wu Mei	43.98	—
Xie Bai	44.33	—
Yuan Zhi	44.70	—
Gan Jiang	44.85	—
Chi Shao (Yao)	46.38	—
(Huai) Niu Xi	46.99	—
Qiang Huo	48.31	—
Du Huo	49.51	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zhi 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
62734	62734	0.00	17.11
62733	62733	0.00	16.57
62529	62529	0.00	16.78
62530	62530	0.00	16.82

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 2.1-2020-01

Validated substance/substance group **Zhu Ling**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 10002254-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G199H1509921	62747	40	from supplier
PhytoComm	Zhu Ling	G199H1509921	62748	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G199H1509921	62747 [†]	20
PhytoComm	Zhu Ling	G199H1509921	62748 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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
Phytocomm	Zhu Ling	G199H1509421	2
PhytoComm	Zhu Ling	G199H1509022	1
PhytoComm	Zhu Ling	G199H1509521	1
Mediherb	Zhu Ling		1

- 848 spectra from 11 *Apo-Ident* customers from a total of 512 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	25 721
Type B	0	40	0	13 127
Type C	0	0	5	848

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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8240 %)	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
Mang Xiao	17.06	—
Tu Fu Ling	24.73	—
Bai He	27.08	—
Mai Ya	29.02	—
E Zhu	31.76	—
San Qi	32.53	—
Ji Li	32.72	—
Sheng Jiang	33.49	—
(Bai) Dou Kou	34.78	—
Sha Shen (Bei)	35.56	—
Jiang Huang	35.72	—
Gui Zhi	35.88	—
Sha Ren	37.36	—
Chuan Xiong	37.96	—
Sang Zhi	39.77	—
Di Gu Pi	39.90	—
San Leng	40.64	—
Mo Yao	40.64	—
Xian He Cao	40.68	—
Jie Geng	41.75	—
Long Dan (Cao)	42.49	—
He Huan Pi	43.83	—
Fu Ling	44.54	—
Yu Jin	47.29	—
Cang Er Zi	48.50	—
Zhu Ru	49.42	—
Ban Xia (Jiang)	49.63	—

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
62747	62747	0.00	17.07
62748	62748	0.00	17.06

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 2.1-2020-01

Validated substance/substance group **Zhu Ru**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60241-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G047HS109QM2	62207	40	from supplier
PhytoComm	Zhu Ru	G047HS109QM2	62208	40	from supplier
PhytoComm	Zhu Ru	G047HS109SK1	62657	40	from supplier
PhytoComm	Zhu Ru	G047HS109SK1	62658	40	from supplier

Validation samples

A total of 39 821 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.
- 25 641 spectra from a total of 317 batches from further 172 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	G047HS109QM2	62207 [†]	20
PhytoComm	Zhu Ru	G047HS109QM2	62208 [†]	20
PhytoComm	Zhu Ru	G047HS109SK1	62657 [†]	20
PhytoComm	Zhu Ru	G047HS109SK1	62658 [†]	20

- 13 087 spectra from a total of 318 batches from further 172 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	2
Phytocomm	Zhu Ru	g047h0625321	1
Phytocomm	Zhu Ru	G047h625321	1
Phytocomm	Zhu Ru	G047H0625422	3
PhytoComm	Zhu Ru	G047H0625321	1

- 844 spectra from 12 *Apo-Ident* customers from a total of 512 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	8	147	13	25 633
Type B	10	63	17	13 077
Type C	0	0	9	844

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	99.9721 % (> 99.9572 %)	91.8750 % (> 90.0000 %)
Type B	99.9322 % (> 99.9033 %)	78.7500 % (> 75.0000 %)
Type C	100.0000 % (> 98.8228 %)	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
Sheng Jiang	2.63	—
Yi Yi Ren	3.30	—
Rou Gui	4.01	—
Fu Xiao Mai	4.63	—
Ban Xia (Jiang)	5.38	—
Gui Zhi	6.16	—
Di Gu Pi	6.77	—
He Huan Pi	7.86	—
Fu Ling	8.12	—
Ci Wu Jia	8.39	—
Ji Li	8.95	—
Bai Bian Dou	9.19	—
Wang Bu Liu Xing	10.47	—
Bai Shao Yao	11.66	—
Lian Zi	12.43	—
Tao Ren	12.78	—
Zhe Bei Mu	13.33	—
Bai Zi Ren	13.35	—
Shen Qu	13.69	—
Ling Zhi	13.84	—
Gou Teng	15.82	—
Huo Ma Ren	16.37	—
Tai Zi Shen	16.49	—
Tian Hua Fen	17.15	—
Gua Lou	17.16	—
Ji Xue Teng	17.57	—
Bai Xian Pi	18.51	—
Fu Shen	19.71	—
Ma Huang Gen	19.92	—
Shan Yao	19.95	—
Yuan Zhi	20.14	—
Pi Pa Ye	21.08	—
Tian Men Dong	22.45	—
Lian Qiao	22.75	—
Yu Jin	23.00	—
Ye Jiao Teng	23.01	—
Ren Dong Teng	23.66	—
Fo Shou	23.76	—
Long Yan Rou	23.77	—
Da Zao	23.80	—
Chen Pi	24.18	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chuan Xiong	25.60	—
Yan Hu Suo	26.18	—
Dang Shen	27.90	—
Tu Fu Ling	28.89	—
Fu Zi	29.25	—
Dan Shen	29.38	—
Shan Yu Rou	29.98	—
Dang Gui Wei	31.11	—
Zhi Mu	31.11	—
Mao Dong Qing	31.27	—
Fang Ji	32.17	—
Wu Wei Zi	32.39	—
Jiao Gu Lan	33.38	—
Dan Dou Chi	33.63	—
San Qi	33.87	—
Jin Yin Hua	34.76	—
Yin Yang Huo	35.99	—
Huang Qin	36.26	—
Che Qian Zi	36.73	—
Zhi Ke	37.16	—
Yu Zhu	37.33	—
Dang Gui	37.75	—
Hou Po	37.81	—
Chai Hu	38.03	—
Mang Xiao	39.44	—
Cang Zhu	40.09	—
Ma Huang	40.92	—
Suan Zao Ren	44.30	—
Huang Lian	47.57	—
Hong Jing Tian	47.59	—
Bai Zhu	47.78	—
Ze Xie	48.57	—
Qiang Huo	48.63	—
Jie Geng	48.66	—
Gou Qi Zi	48.67	—
Sang Zhi	49.52	—
Qing Pi	49.70	—

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	3.31
62208	62208	0.00	2.82
62207	62207	0.00	2.63
62658	62658	0.00	3.30

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 2.1-2020-01

Validated substance/substance group **Zi Hua Di Ding**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 50376-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Zi Hua Di Ding; Viola 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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Zi Hua Di Ding	1	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

Validation samples

A total of 39 821 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 Hua Di Ding*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 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

- 13 127 spectra from a total of 319 batches from further 172 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

- 851 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 *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	80	0	25 720
Type B	1	39	1	13 126
Type C	0	0	2	851

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 % (> 99.9815 %)	100.0000 % (> 92.5000 %)
Type B	99.9927 % (> 99.9640 %)	97.5000 % (> 90.0000 %)
Type C	100.0000 % (> 98.8279 %)	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 *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
Dan Shen	4.69	—
Mu Dan Pi	5.46	—
Qing Pi	6.76	—
Jiao Gu Lan	8.16	—
Hong Jing Tian	8.67	—
Huang Bai	9.37	—
Fang Ji	9.38	—
Hou Po	9.67	—
Chai Hu	10.95	—
Jin Yin Hua	11.06	—
Qiang Huo	11.32	—
Yin Yang Huo	12.18	—
Chuan Xiong	14.27	—
Shan Yao	14.42	—
Pi Pa Ye	14.49	—
Zhi Ke	15.09	—
Wu Wei Zi	15.15	—
Ye Jiao Teng	15.17	—
Mao Dong Qing	15.39	—
Du Zhong	16.69	—
Ma Huang	17.01	—
Dan Dou Chi	17.01	—
Yan Hu Suo	17.12	—
Ren Dong Teng	17.23	—
Bai Xian Pi	17.35	—
Bai Bian Dou	17.36	—
Huang Lian	17.67	—
Gua Lou	17.74	—
Bai Shao Yao	18.59	—
Fu Zi	18.91	—
Yu Jin	19.63	—
Ji Xue Teng	19.81	—
Cang Zhu	19.91	—
Ji Li	19.96	—
Sang Ye	20.17	—
Shan Yu Rou	20.74	—
Yuan Zhi	21.45	—
Bai Zhu	21.64	—
Che Qian Zi	22.09	—
Lian Qiao	22.31	—
Gou Qi Zi	22.44	—
Lian Zi	23.42	—
Di Gu Pi	23.67	—
He Huan Pi	24.03	—
Shen Qu	24.19	—
Ling Zhi	24.26	—
Hong Hua	24.52	—
Tu Fu Ling	27.04	—
Cang Er Zi	27.13	—
Gui Zhi	27.44	—
Pu Gong Ying	27.99	—
Gou Teng	29.03	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Chen Pi	29.54	—
Tao Ren	29.75	—
Zhe Bei Mu	31.36	—
San Qi	31.42	—
Xi Xian Cao	31.73	—
Huo Ma Ren	32.02	—
Rou Gui	32.11	—
Suan Zao Ren	32.16	—
Tian Hua Fen	33.16	—
Da Zao	33.27	—
Huang Qin	33.79	—
Chi Shao (Yao)	34.14	—
Gan Cao	34.30	—
Bai Zi Ren	34.36	—
Dang Gui	34.47	—
Ze Xie	34.71	—
Dang Shen	35.08	—
Tai Zi Shen	35.36	—
Ban Xia (Jiang)	35.46	—
Sheng Jiang	35.67	—
Bai Mao Gen	36.26	—
Yi Mu Cao	36.32	—
Chuang Mu Xiang	36.35	—
Long Yan Rou	36.72	—
Wu Mei	36.99	—
Sang Zhi	37.00	—
Ma Huang Gen	37.48	—
Mang Xiao	37.58	—
Jie Geng	37.73	—
Zhi Mu	37.77	—
Fu Xiao Mai	37.93	—
Xia Ku Cao	38.53	—
Ban Zhi Lian	38.87	—
Fu Ling	38.95	—
Dang Gui Wei	39.65	—
Jiang Huang	39.74	—
Yu Zhu	40.20	—
Ci Wu Jia	40.97	—
San Leng	41.42	—
Zhu Ru	45.38	—
Yi Yi Ren	45.42	—
Zi Wan	45.65	—
Fu Shen	45.71	—
Wu Yao	46.08	—
(Shi) Chang Pu	46.38	—
Chuan Lian Zi	47.33	—
E Zhu	47.49	—
Long Dan (Cao)	49.65	—

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
62628	62628	0.00	4.69
62627	62627	0.00	6.76

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 2.1-2020-01

Validated substance/substance group **Zi Su Zi**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60133-2020-03-12
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

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	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	G183H1213822	62483	40	from supplier
PhytoComm	Zi Su Zi	G183H1213822	62484	40	from supplier

Validation samples

A total of 39 821 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.
- 25 721 spectra from a total of 318 batches from further 172 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	G183H1213822	62483 [†]	20
PhytoComm	Zi Su Zi	G183H1213822	62484 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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.
- 853 spectra from 12 *Apo-Ident* customers from a total of 516 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	25 721
Type B	0	40	0	13 127
Type C	0	0	0	853

The substance/substance group *Zi Su 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	100.0000 % (> 98.8323 %)	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
Cang Er Zi	9.52	—
Huo Ma Ren	12.35	—
Gan Jiang	16.24	—
San Qi	17.99	—
Xie Bai	18.37	—
Bai Bian Dou	19.18	—
Xiang Fu	22.48	—
Gua Lou	23.30	—
Gou Qi Zi	23.40	—
Ji Li	23.68	—
Sha Shen (Bei)	24.41	—
She Gan	24.46	—
Jiang Huang	24.58	—
Chai Hu	24.66	—
Di Gu Pi	24.73	—
Mang Xiao	24.95	—
Sang Zhi	25.08	—
Zhi Mu	25.45	—
E Zhu	25.58	—
Niu Bang Zi	26.48	—
Shan Yao	27.29	—
Tian Hua Fen	27.58	—
Gui Zhi	27.89	—
Yan Hu Suo	28.24	—
Chen Pi	28.53	—
Yi Mu Cao	29.18	—
Jie Geng	29.34	—
Dang Shen	29.66	—
Long Dan (Cao)	29.94	—
Chuan Lian Zi	29.98	—
Sheng Jiang	30.33	—
Hou Po	30.45	—
Wang Bu Liu Xing	30.65	—
Sha Ren	30.71	—
Ba Ji Tian	30.76	—
Mu Gua	31.07	—
Fang Feng	31.08	—
(Huai) Niu Xi	31.09	—
(Bai) Dou Kou	31.40	—
Cang Zhu	32.10	—
(Shi) Chang Pu	32.44	—
Chuan Xiong	32.90	—
Mo Yao	32.97	—
Bai He	35.47	—
Dong Gua Zi	35.59	—
Dang Gui Wei	35.76	—
Lai Fu Zi	36.08	—
Dang Gui	36.37	—
Ku Shen	36.52	—
Qing Pi	36.63	—
Du Huo	37.67	—
Huang Qi	37.86	—
Ban Xia (Jiang)	37.98	—
Zi Wan	38.62	—
Qiang Huo	38.65	—
Wei Ling Xian	40.33	—
Suan Zao Ren	40.84	—
He Huan Pi	41.11	—
Mai Men Dong	41.19	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Bai Zhu	41.54	—
Lu Gen	41.78	—
Yuan Zhi	42.01	—
Huang Lian	42.24	—
Bai Mao Gen	42.32	—
Mai Ya	42.67	—
Wu Yao	42.75	—
Fu Ling	43.10	—
Fu Zi	43.31	—
Nü Zhen Zi	43.43	—
Lian Qiao	43.76	—
Xian He Cao	43.85	—
Zhi Zi	44.02	—
Xu Duan	44.06	—
Yu Jin	44.20	—
Yin Yang Huo	44.79	—
Ze Xie	44.83	—
Pu Gong Ying	45.95	—
Huang Bo	46.02	—
Bai Xian Pi	46.85	—
Jin Yin Hua	46.89	—
Fang Ji	47.20	—
Wu Wei Zi	47.94	—
Lian Zi	48.04	—
Mu Dan Pi	48.22	—
Sang Bai Pi	48.75	—
San Leng	48.85	—
Ma Bian Cao	48.91	—
Tu Fu Ling	49.41	—

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
62484	62484	0.00	9.52
62483	62483	0.00	9.60

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 2.1-2020-01

Validated substance/substance group **Zi Wan**
Substance class TCM - Granulated herbal extracts (PhytoComm)
Report date 12/03/2020
Report number 60155-2020-03-12
Executing company HiperScan GmbH
Weißeritzstraße 3
01067 Dresden
Germany

Relevant substance names

Zi Wan; Asteris tatarici radix

Special notes

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

no information

Applicable documents

978-3-7692-6679-5 *European Pharmacopoeia 9th Edition, Basic Version 2017* [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	Type A	Type B	Type C
Zi Wan	1	0	1

Second-stage model

For differentiation of the substance/substance group *Zi Wan* 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 Wan*:

Supplier	Substance	Batch	Sample ID	Spectra	Certificate
PhytoComm	Zi Wan	G038H1215621	62045	40	from supplier
PhytoComm	Zi Wan	G038H1215621	62046	40	from supplier

Validation samples

A total of 39 821 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 Wan*. These samples are listed above in the [calibration samples](#) section. The reference samples come from 1 different batches.
- 25 721 spectra from a total of 318 batches from further 172 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 Wan*.
- 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 Wan	G038H1215621	62045 [†]	20
PhytoComm	Zi Wan	G038H1215621	62046 [†]	20

- 13 127 spectra from a total of 319 batches from further 172 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 *Zi Wan*.
- 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 Wan	G038H1215321	2

- 851 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 *Zi Wan* 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 Wan* 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	25 721
Type B	0	40	0	13 127
Type C	9	2	0	842

The substance/substance group *Zi Wan* 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 % (> 99.9703 %)	100.0000 % (> 92.5000 %)
Type B	100.0000 % (> 99.9424 %)	100.0000 % (> 85.0000 %)
Type C	99.0310 % (> 98.4449 %)	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 *Zi Wan* 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 Xie	15.27	—
Lian Qiao	15.29	—
Chi Shao (Yao)	16.82	—
Yuan Zhi	18.05	—
She Gan	19.80	—
Dang Gui Wei	20.72	—
(Shi) Chang Pu	22.30	—
Huang Qi	22.76	—
Xiang Fu	23.68	—
Gua Lou	24.57	—
Du Huo	25.11	—
Wu Yao	25.74	—
Qing Pi	25.84	—
Bai Mao Gen	26.00	—
Wei Ling Xian	26.38	—
(Huai) Niu Xi	26.43	—
Zhi Mu	26.88	—
Yi Mu Cao	26.98	—
Sang Zhi	27.11	—
Ju Hua	27.81	—
Gou Qi Zi	28.04	—
Ku Shen	28.41	—
Cang Zhu	28.84	—
Niu Bang Zi	30.01	—
Hong Jing Tian	30.42	—
San Leng	30.81	—
Chai Hu	30.95	—
Mu Dan Pi	32.14	—
Cang Er Zi	32.57	—
Gan Cao	32.87	—
Gui Zhi	33.01	—
Chuan Xiong	33.46	—
Di Gu Pi	33.54	—
Jie Geng	33.57	—
Zhi Zi	34.81	—
Zhi Ke	35.62	—
Dang Gui	35.63	—
Fang Feng	36.67	—
Mang Xiao	36.92	—
Mu Gua	38.06	—
Jin Yin Hua	38.14	—
Wu Wei Zi	38.65	—
Shan Yao	39.64	—
Huo Ma Ren	39.96	—
Lu Gen	40.02	—
Gan Jiang	40.24	—
Xu Duan	42.02	—
Huang Bo	42.53	—
Shan Yu Rou	42.59	—
Pu Gong Ying	42.93	—
Dang Shen	44.99	—
Fu Zi	45.02	—

continued on the next page

continued from previous page

Substanz	Distance in main model	Distance in second-stage model
Huang Lian	45.18	—
Jiang Huang	45.25	—
Xie Bai	45.70	—
Chen Pi	45.92	—
Hou Po	46.30	—
Ji Li	46.82	—
Wu Zhu Yu	47.58	—
Dan Shen	47.62	—
Mao Dong Qing	48.87	—
Bai Shao Yao	48.87	—
Lian Zi	49.46	—
Sha Ren	49.79	—
Bai Zhu	49.80	—
Lai Fu Zi	49.90	—
Xin Yi	50.02	—

The list stops after the first substance with a *Mahalanobis distance* greater than 50. If the substance/substance group *Zi Wan* 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
62045	62045	0.00	15.27
62046	62046	0.00	15.40

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.

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 108 batches. From these, 145 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
PhytoComm	(Fen) Bi Xie	G0290H1251121	1
Phytocomm	Bai Bu	G235H0608321	1
PhytoComm	Bai Bu	G235H0608221	1
Phytocomm	Bai Guo	G114H0509321	2
Phytocomm	Bai Guo	G114H0509521	1
Phytocomm	Bai Jiang Cao	G181H1180321	1
Phytocomm	Bai Jiang Cao	G181H1180422	1
Phytocomm	Bai Tou Weng	G207H0511521	4
PhytoComm	Bai Tou Weng	G207H0511221	3
Phytocomm	Bai Zhi	G020H0503421	1
Phytocomm	Ban Lan Gen	G128H0816421	2
PhytoComm	Bie Jia	G507H2401121	1
PhytoComm	Bie Jia	G507H2401321	1
Phytocomm	Bo He	g158h1712022	1
PhytoComm	Bo He	G158H1712121	1
PhytoComm	Cao Guo	E241050	1
phytocomm	Ce Bai Ye	g050h0975321	2
Phytocomm	Ce Bai Ye	G050H0975321	1
Phytocomm	Ce Bai Ye	G050HS245	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
Phytocomm	Chuan Mu Tong	g009h0310321	1
Phytocomm	Chuan Mu Tong	G009H0310321	1
PhytoComm	Chuan Mu Tong	G009H0310521	1
Phytocomm	Chuan Niu Xi	g318h0431321	1
PhytoComm	Chuan Niu Xi	G318H0431321	1
Phytocomm	Da Fu Pi	G293H0321221	1
Phytocomm	Da Qing Ye	G127H0318421	2
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
Phytocomm	Ding Xiang	G058H0210021	1
PhytoComm	E Bu Shi Cao	G061H180122	1
Phytocomm	E Jiao	g036h0861121	1

continued on the next page

continued from previous page

Supplier	Substance	Batch	Spectra
Phytocomm	E Jiao	g036h0861221	1
Phytocomm	E Jiao	G036H0861421	2
PhytoComm	E Jiao	G036H0861221	4
Phytocomm	Fu Pen Zi	g213h1813122	1
PhytoComm	Geng Mi	G329H1320121	1
Phytocomm	Gua Lou Ren	g243h1095123	1
PhytoComm	Gua Lou Ren	G243H1095421	1
PhytoComm	He Zi	G291H1239221	2
Phytocomm	Huai Hua	G263H1423421	1
Phytocomm	Ji Nei Jin	G259H2102321	1
Phytocomm	Ji Nei Jin	g259h2102321	1
Phytocomm	Jie Cao	H2111921	1
PhytoComm	Kuan Dong Hua	H1219921	1
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
PhytoComm	Ma Chi Xian	G255F110310	1
Phytocomm	Mi Huan Jun	G510H0418221	1
Phytocomm	Mi Huan Jun	G510H0418422	1
Phytocomm	Mu Zei	G265H0427421	1
PhytoComm	Mu Zei	G265H0427823	1
Phytocomm	Qian Hu	G185H0926322	4
Phytocomm	Qian Shi	G102H0836422	1
PhytoComm	Qian Shi	G102H0836121	1
Phytocomm	Quan Xie	G222F110314	1
Phytocomm	Ren Shen	g115ho562121	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	Suo Yang	G260H1831121	1
Phytocomm	Wu Gong	G311H1411321	1
Phytocomm	Wu Gong	G311H411321	1
Phytocomm	Xi Yang Shen	G324H0647321	1
Phytocomm	Xi Yang Shen	g324h0647321	1
Phytocomm	Xi Yang Shen	G324HS4070L1	1
Phytocomm	Xian Mao	G304H0549321	2
Phytocomm	Xian Mao	G304H0549521	2
Phytocomm	Xiao Hui Xiang	g104h0339321	1
Phytocomm	Xiao Hui Xiang	G104H0339522	3
Phytocomm	Xuan Fu Hua	G126H1168521	1
Phytocomm	Yi Zhi Ren	G015H1051321	2
Phytocomm	Yi Zhi Ren	G015H1051521	1
Phytocomm	Ze Lan	g268h1606221	1
Phytocomm	Ze Lan	G268H1606421	1
PhytoComm	Ze Lan	G268H1606221	2
Phytocomm	Zhen Zhu Mu	g155h0979221	1
Phytocomm	Zhi Gan Cao	g118h0881322	1
Phytocomm	Zhi Gan Cao	G118H0881522	4
PhytoComm	Zhi Gan Cao	G118H0881322	1
PhytoComm	Zi Su Ye	G182H1214321	3

continued on the next page

continued from previous page

Supplier	Substance	Batch	Spectra
PhytoComm	Zi Su Ye	G182HS2680P1	1

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).

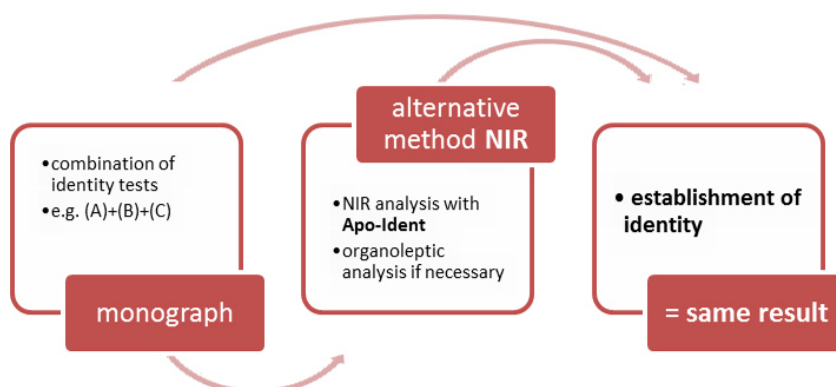


Figure 2: The combination of tests of the monograph is replaced by the alternative method NIR spectroscopy using *Apo-Ident*. This is permissible because both test procedures result in the establishment of the identity of the raw material.

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 2017*. 9. Ausgabe. Deutscher Apotheker Verlag (978-3-7692-6679-5)
- [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] JOVANOVIĆ, 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 ; RÖTTA: Apothekenbetriebsordnung, Kommentar § 6 Rn. 10, 2010

Index

- (Bai) Dou Kou, [21](#)
(Huai) Niu Xi, [27](#)
(Ku) Xing Ren, [33](#)
(Sheng) Di Huang, [37](#)
(Shi) Chang Pu, [43](#)
- Abutili semen, *see* Dong Kui Zi, [253](#)
Acanthopanax cortex, *see* Wu Jia Pi, [841](#)
Achyranthis bidentatae radix, *see* (Huai) Niu Xi, [27](#)
Aconiti radix lateralis praep., *see* Fu Zi, [311](#)
Acori rhizoma, *see* (Shi) Chang Pu, [43](#)
Agrimoniae herba, *see* Xian He Cao, [881](#)
Ai Ye, [49](#)
Albiziae cortex, *see* He Huan Pi, [381](#)
Alismatis rhizoma, *see* Ze Xie, [977](#)
Allii bulbosus, *see* Xie Bai, [893](#)
Amomi cardamomi fructus, *see* (Bai) Dou Kou, [21](#)
Amomi villosi fructus, *see* Sha Ren, [737](#)
Anemarrhenae rhizoma, *see* Zhi Mu, [995](#)
Angelicae pubescentis radix, *see* Du Huo, [257](#)
Angelicae sinensis radix, *see* Dang Gui, [221](#)
Angelicae sinensis radix extremitas, *see* Dang Gui Wei, [227](#)
Arctii fructus, *see* Niu Bang Zi, [633](#)
Armeniacae amarum semen, *see* (Ku) Xing Ren, [33](#)
Artemisiae apiacae herba, *see* Qing Hao, [669](#)
Artemisiae argyi folium, *see* Ai Ye, [49](#)
Artemisiae scopariae herba, *see* Yin Chen Hao, [943](#)
Asparagi cochinchinensis radix, *see* Tian Men Dong, [811](#)
Asteris tatarici radix, *see* Zi Wan, [1037](#)
Astragali membranacei radix, *see* Huang Qi, [445](#)
Atractylodis macrocephalae rhizoma, *see* Bai Zhu, [95](#)
Atractylodis rhizoma, *see* Cang Zhu, [131](#)
Aurantii fructus immaturus, *see* Zhi Shi, [1001](#)
- Ba Ji Tian, [53](#)
Bai Bian Dou, [59](#)
Bai He, [65](#)
Bai Hua She She Cao, [71](#)
Bai Mao Gen, [77](#)
Bai Shao Yao, [83](#)
Bai Xian Pi, [89](#)
Bai Zhu, [95](#)
Bai Zi Ren, [101](#)
Bambusae in taeniae caulis, *see* Zhu Ru, [1019](#)
Ban Xia (Jiang), [107](#)
Ban Zhi Lian, [113](#)
Belamcandae rhizoma, *see* She Gan, [761](#)
Benincasae hispidae semen, *see* Dong Gua Zi, [249](#)
Biotae orientalis semen, *see* Bai Zi Ren, [101](#)
Bu Gu Zhi, [119](#)
Bupleuri radix, *see* Chai Hu, [137](#)
- Cang Er Zi, [125](#)
Cang Zhu, [131](#)
Cannabis semen, *see* Huo Ma Ren, [457](#)
Carthami flos, *see* Hong Hua, [393](#)
Cassiae torae semen, *see* Jue Ming Zi, [519](#)
Chaenomelis lagenariae fructus, *see* Mu Gua, [627](#)
Chai Hu, [137](#)
Che Qian Zi, [145](#)
Chen Pi, [151](#)
Chi Shao (Yao), [159](#)
Chrysanthemi flos, *see* Ju Hua, [513](#)
Chuan Lian Zi, [165](#)
Chuan Xiong, [171](#)
Chuang Mu Xiang, [179](#)
Ci Wu Jia, [185](#)
Cinnamomi cassiae cortex, *see* Rou Gui, [691](#)
Cinnamomi cassiae ramulus, *see* Gui Zhi, [367](#)
Cistanchis herba, *see* Rou Cong Rong, [685](#)
Citri aurantii fructus, *see* Zhi Ke, [989](#)
Citri reticulatae pericarpium, *see* Chen Pi, [151](#)
Citri reticulatae viride pericarpium, *see* Qing Pi, [673](#)
Citri sarcodactylis fructus, *see* Fo Shou, [287](#)
Clematidis radix, *see* Wei Ling Xian, [835](#)
Codonopsis pilosulae radix, *see* Dang Shen, [233](#)
Coicis semen, *see* Yi Yi Ren, [937](#)
Coptidis rhizoma, *see* Huang Lian, [439](#)
Corni officinalis fructus, *see* Shan Yu Rou, [755](#)
Corydalis rhizoma, *see* Yan Hu Suo, [917](#)
Curcumae longae rhizoma, *see* Jiang Huang, [477](#)
Curcumae tuber, *see* Yu Jin, [955](#)
Curcumae zedoariae rhizoma, *see* E Zhu, [269](#)
Cuscutae semen, *see* Tu Si Zi, [825](#)
Cyperis rotundi rhizoma, *see* Xiang Fu, [887](#)
- Da Huang, [191](#)
Da Zao, [195](#)
Dan Dou Chi, [201](#)
Dan Shen, [207](#)
Dan Zhu Ye, [215](#)
Dang Gui, [221](#)
Dang Gui Wei, [227](#)
Dang Shen, [233](#)
Di Fu Zi, [239](#)

- Di Gu Pi, [243](#)
Dictamni cortex, *see* Bai Xian Pi, [89](#)
Dioscoreae oppositae rhizoma, *see* Shan Yao, [749](#)
Dipsaci radix, *see* Xu Duan, [905](#)
Dolichoris lablab semen, *see* Bai Bian Dou, [59](#)
Dong Gua Zi, [249](#)
Dong Kui Zi, [253](#)
Drynariae rhizoma, *see* Gu Sui Bu, [349](#)
Du Huo, [257](#)
Du Zhong, [263](#)

E Zhu, [269](#)
Ecliptae prostratae herba, *see* Han Lian Cao, [375](#)
Eleutherococci radix, *see* Ci Wu Jia, [185](#)
Ephedrae herba, *see* Ma Huang, [577](#)
Ephedrae radix, *see* Ma Huang Gen, [583](#)
Epimedii herba, *see* Yin Yang Huo, [949](#)
Eriobotryae japonicae folium, *see* Pi Pa Ye, [645](#)
Eucommiae ulmoidis cortex, *see* Du Zhong, [263](#)
Euphoriae longanae arillus, *see* Long Yan Rou, [559](#)
Evodiae rutecarpae fructus, *see* Wu Zhu Yu, [863](#)

Fang Feng, [275](#)
Fang Ji, [281](#)
Fo Shou, [287](#)
Forsythiae fructus, *see* Lian Qiao, [535](#)
Fritillariae thunbergii bulbus, *see* Zhe Bei Mu, [983](#)
Fu Ling, [293](#)
Fu Shen, [299](#)
Fu Xiao Mai, [305](#)
Fu Zi, [311](#)

Gan Cao, [319](#)
Gan Jiang, [325](#)
Ganoderma, *see* Ling Zhi, [547](#)
Gardeniae fructus, *see* Zhi Zi, [1007](#)
Ge Gen, [331](#)
Gentianae macrophyllae radix, *see* Qin Jiao, [663](#)
Gentianae scabrae radix, *see* Long Dan (Cao), [553](#)
Glehniae radix, *see* Sha Shen (Bei), [743](#)
Glycyrrhizae radix, *see* Gan Cao, [319](#)
Gou Qi Zi, [337](#)
Gou Teng, [343](#)
Gu Sui Bu, [349](#)
Gua Lou, [355](#)
Guang Huo Xiang, [363](#)
Gui Zhi, [367](#)
Gynostemma herba, *see* Jiao Gu Lan, [483](#)
Gypsum fibrosum, *see* Shi Gao, [779](#)

Han Lian Cao, [375](#)
He Huan Pi, [381](#)
He Shou Wu, [387](#)
Hong Hua, [393](#)
Hong Jing Tian, [399](#)
Hordei vulgaris fructus germinatus, *see* Mai Ya, [595](#)
Hou Po, [405](#)
Houttuyniae herba, *see* Yu Xing Cao, [961](#)
Hu Zhang, [411](#)
Hua Shi, [417](#)
Huang Bai, [421](#)
Huang Bo, [427](#)
Huang Jing, [433](#)
Huang Lian, [439](#)
Huang Qi, [445](#)
Huang Qin, [451](#)
Huo Ma Ren, [457](#)

Ilicis pubescendis radix, *see* Mao Dong Qing, [609](#)
Imperatae cylindrica rhizoma, *see* Bai Mao Gen, [77](#)

Ji Li, [463](#)
Ji Xue Teng, [471](#)
Jiang Huang, [477](#)
Jiao Gu Lan, [483](#)
Jie Geng, [489](#)
Jin Qian Cao, [495](#)
Jin Yin Hua, [501](#)
Jing Jie, [507](#)
Ju Hua, [513](#)
Jue Ming Zi, [519](#)

Kochiae scopariae fructus, *see* Di Fu Zi, [239](#)
Ku Shen, [523](#)

Lai Fu Zi, [529](#)
Leonuri heterophylli herba, *see* Yi Mu Cao, [931](#)
Lian Qiao, [535](#)
Lian Zi, [541](#)
Ligustici wallichii radix, *see* Chuan Xiong, [171](#)
Ligustri lucidi fructus, *see* Nü Zhen Zi, [639](#)
Lilii bulbus, *see* Bai He, [65](#)
Linderae radix, *see* Wu Yao, [857](#)
Ling Zhi, [547](#)
Long Dan (Cao), [553](#)
Long Yan Rou, [559](#)
Lonicerae japonicae caulis, *see* Ren Dong Teng, [679](#)
Lonicerae japonicae flos, *see* Jin Yin Hua, [501](#)
Lophatheri gracilis herba, *see* Dan Zhu Ye, [215](#)
Lu Gen, [565](#)
Lycii chinensis fructus, *see* Gou Qi Zi, [337](#)

- Lycii chinensis radicis cortex, *see* Di Gu Pi, [243](#)
- Lysimachiae herba, *see* Jin Qian Cao, [495](#)
- Ma Bian Cao, [571](#)
- Ma Huang, [577](#)
- Ma Huang Gen, [583](#)
- Magnoliae flos, *see* Xin Yi, [899](#)
- Magnoliae officinalis cortex, *see* Hou Po, [405](#)
- Mai Men Dong, [589](#)
- Mai Ya, [595](#)
- Man Jing Zi, [601](#)
- Mang Xiao, [605](#)
- Mao Dong Qing, [609](#)
- Massa fermentata medicinalis, *see* Shen Qu, [767](#)
- Meliae toosendan fructus, *see* Chuan Lian Zi, [165](#)
- Mirabilitum, *see* Mang Xiao, [605](#)
- Mo Yao, [615](#)
- Mori albae ramulus, *see* Sang Zhi, [731](#)
- Mori albi folium, *see* Sang Ye, [725](#)
- Mori cortex, *see* Sang Bai Pi, [713](#)
- Morindae officinalis radix, *see* Ba Ji Tian, [53](#)
- Moutan cortex radicis, *see* Mu Dan Pi, [621](#)
- Mu Dan Pi, [621](#)
- Mu Gua, [627](#)
- Myrrha resina, *see* Mo Yao, [615](#)
- Nü Zhen Zi, [639](#)
- Nelumbinis semen, *see* Lian Zi, [541](#)
- Niu Bang Zi, [633](#)
- Notoginseng radix, *see* San Qi, [707](#)
- Notopterygii rhizoma et radix, *see* Qiang Huo, [657](#)
- Oldenlandiae diffusae herba, *see* Bai Hua She She Cao, [71](#)
- Olibanum, *see* Ru Xiang, [697](#)
- Ophiopogonis radix, *see* Mai Men Dong, [589](#)
- Paeoniae lactiflorae albus radix, *see* Bai Shao Yao, [83](#)
- Paeoniae radix rubra, *see* Chi Shao (Yao), [159](#)
- Perillae fructus, *see* Zi Su Zi, [1031](#)
- Persicae semen, *see* Tao Ren, [799](#)
- Phellodendri chinensis cortex, *see* Huang Bai, [421](#)
- Phellodendri cortex chinensis, *see* Huang Bo, [427](#)
- Phragmitis communis rhizoma, *see* Lu Gen, [565](#)
- Pi Pa Ye, [645](#)
- Pinelliae rhizoma praeparatum cum zingibere, *see* Ban Xia (Jiang), [107](#)
- Plantaginis semen, *see* Che Qian Zi, [145](#)
- Platycodi grandiflori radix, *see* Jie Geng, [489](#)
- Pogostemonis herba, *see* Guang Huo Xiang, [363](#)
- Polygalae radix, *see* Yuan Zhi, [971](#)
- Polygonati odorati rhizoma, *see* Yu Zhu, [965](#)
- Polygonati rhizoma, *see* Huang Jing, [433](#)
- Polygoni cuspidati rhizoma et radix, *see* Hu Zhang, [411](#)
- Polygoni multiflori caulis, *see* Ye Jiao Teng, [925](#)
- Polygoni multiflori radix, *see* He Shou Wu, [387](#)
- Polypori umbellati sclerotium, *see* Zhu Ling, [1013](#)
- Poriae cocos sclerotium, *see* Fu Ling, [293](#)
- Poriae cocos sclerotium paradidicis, *see* Fu Shen, [299](#)
- Prunellae vulgaris spica, *see* Xia Ku Cao, [875](#)
- Pruni mume fructus, *see* Wu Mei, [845](#)
- Pseudostellariae heterophyllae radix, *see* Tai Zi Shen, [793](#)
- Psoraleae corylifoliae fructus, *see* Bu Gu Zhi, [119](#)
- Pu Gong Ying, [651](#)
- Puerariae radix, *see* Ge Gen, [331](#)
- Qiang Huo, [657](#)
- Qin Jiao, [663](#)
- Qing Hao, [669](#)
- Qing Pi, [673](#)
- Raphani sativi semen, *see* Lai Fu Zi, [529](#)
- Rehmanniae radix, *see* (Sheng) Di Huang, [37](#)
- Rehmanniae radix praep., *see* Shu Di (Huang), [783](#)
- Ren Dong Teng, [679](#)
- Rhei radix et rhizoma, *see* Da Huang, [191](#)
- Rhodiolae crenulatae radix, *see* Hong Jing Tian, [399](#)
- Rou Cong Rong, [685](#)
- Rou Gui, [691](#)
- Ru Xiang, [697](#)
- Salviae miltiorrhizae radix, *see* Dan Shen, [207](#)
- San Leng, [701](#)
- San Qi, [707](#)
- Sang Bai Pi, [713](#)
- Sang Ji Shend, [719](#)
- Sang Ye, [725](#)
- Sang Zhi, [731](#)
- Saposhnikoviae radix, *see* Fang Feng, [275](#)
- Schisandrae chinensis fructus, *see* Wu Wei Zi, [851](#)
- Schizonepetae tenuifoliae herba, *see* Jing Jie, [507](#)
- Scrophulariae ningpoensis radix, *see* Xuan Shen, [911](#)
- Scutellariae baicalensis radix, *see* Huang Qin, [451](#)

- Scutellariae barbatae herba, *see* Ban Zhi Lian, [113](#)
- Sha Ren, [737](#)
- Sha Shen (Bei), [743](#)
- Shan Yao, [749](#)
- Shan Yu Rou, [755](#)
- She Gan, [761](#)
- Shen Qu, [767](#)
- Sheng Jiang, [773](#)
- Shi Gao, [779](#)
- Shu Di (Huang), [783](#)
- Siegesbeckiae herba, *see* Xi Xian Cao, [869](#)
- Smilacis glabrae rhizoma, *see* Tu Fu Ling, [817](#)
- Sojae semen praeparatum, *see* Dan Dou Chi, [201](#)
- Sophorae flavescentis radix, *see* Ku Shen, [523](#)
- Sparganii rhizoma, *see* San Leng, [701](#)
- Spatholobi caulis, *see* Ji Xue Teng, [471](#)
- Stephaniae tetrandrae radix, *see* Fang Ji, [281](#)
- Suan Zao Ren, [787](#)
- Tai Zi Shen, [793](#)
- Talcum, *see* Hua Shi, [417](#)
- Tao Ren, [799](#)
- Taraxaci mongolici herba cum radice, *see* Pu Gong Ying, [651](#)
- Taxilli herba, *see* Sang Ji Shend, [719](#)
- Tian Hua Fen, [805](#)
- Tian Men Dong, [811](#)
- Tribuli terrestris fructus, *see* Ji Li, [463](#)
- Trichosanthis fructus, *see* Gua Lou, [355](#)
- Trichosanthis radix, *see* Tian Hua Fen, [805](#)
- Tritici aestivi semen levis, *see* Fu Xiao Mai, [305](#)
- Tu Fu Ling, [817](#)
- Tu Si Zi, [825](#)
- Uncariae ramulus cum uncis, *see* Gou Teng, [343](#)
- Vaccariae semen, *see* Wang Bu Liu Xing, [829](#)
- Verbenae officinalis herba, *see* Ma Bian Cao, [571](#)
- Violae herba, *see* Zi Hua Di Ding, [1025](#)
- Viticis fructus, *see* Man Jing Zi, [601](#)
- Vladimirae radix, *see* Chuang Mu Xiang, [179](#)
- Wang Bu Liu Xing, [829](#)
- Wei Ling Xian, [835](#)
- Wu Jia Pi, [841](#)
- Wu Mei, [845](#)
- Wu Wei Zi, [851](#)
- Wu Yao, [857](#)
- Wu Zhu Yu, [863](#)
- Xanthii fructus, *see* Cang Er Zi, [125](#)
- Xi Xian Cao, [869](#)
- Xia Ku Cao, [875](#)
- Xian He Cao, [881](#)
- Xiang Fu, [887](#)
- Xie Bai, [893](#)
- Xin Yi, [899](#)
- Xu Duan, [905](#)
- Xuan Shen, [911](#)
- Yan Hu Suo, [917](#)
- Ye Jiao Teng, [925](#)
- Yi Mu Cao, [931](#)
- Yi Yi Ren, [937](#)
- Yin Chen Hao, [943](#)
- Yin Yang Huo, [949](#)
- Yu Jin, [955](#)
- Yu Xing Cao, [961](#)
- Yu Zhu, [965](#)
- Yuan Zhi, [971](#)
- Ze Xie, [977](#)
- Zhe Bei Mu, [983](#)
- Zhi Ke, [989](#)
- Zhi Mu, [995](#)
- Zhi Shi, [1001](#)
- Zhi Zi, [1007](#)
- Zhu Ling, [1013](#)
- Zhu Ru, [1019](#)
- Zi Hua Di Ding, [1025](#)
- Zi Su Zi, [1031](#)
- Zi Wan, [1037](#)
- Zingiberis rhizoma, *see* Gan Jiang, [325](#)
- Zingiberis rhizoma recens, *see* Sheng Jiang, [773](#)
- Zizyphi jujubae fructus, *see* Da Zao, [195](#)
- Zizyphi spinosae semen, *see* Suan Zao Ren, [787](#)