EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

27 January 2014
EMEA/CHMP/CVMP/QWP/17760/2009 Rev2
Committee for Human Medicinal Products (CHMP)
Committee for Veterinary Medicinal Products (CVMP)

Guideline on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations

Draft Agreed by CHMP / CVMP Quality Working Party	December 2011
Adoption by CHMP for release for consultation	January 2012
Adoption by CVMP for release for consultation	January 2012
End of consultation (deadline for comments)	30 April 2012
Agreed by QWP	January 2014
Adoption by CHMP	February 2014
Adoption by CVMP	March 2014
Date for coming into effect	6 months after publication

This guideline replaces the note for guidance on the use of near infrared spectroscopy by the pharmaceutical industry and the data requirements for new submissions and variations, CPMP/QWP/3309/01 and EMEA/CVMP/961/01

Keywords	***NIR, NIRS, PAT, near infrared Spectroscopy, Process Analytical Technology***

7 Westferry Circus • Canary Wharf • London E14 4HB • United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu

Table of contents

Executive summary **4**

1. Introduction **4**

2. Scope **6**

3. Legal basis **6**

4. General requirements **6**

4.1. Development 7
4.1.1. NIRS procedure scope 7
4.1.2. Description of the NIRS method 8
4.1.3. Description of the NIRS model 8
4.1.4. Feasibility study 8
4.1.5. Risk assessment of variables affecting NIRS procedures 9
4.2. Data collection 10
4.2.1. Sample preparation and presentation 10
4.2.2. Sample population 10
4.2.3. Pre-treatment of data 11
4.2.4. Analysis by the reference analytical procedure 11
4.2.5. Establishment of a spectral library 12
4.3. Data interpretation 12
4.3.1. Statistical spectral quality test 12
4.3.2. Outliers in sample data 13
4.3.3. Out Of Specification (OOS) results in routine batch analysis 13
4.4. Calibration 13
4.5. Validation 14
4.6. Process Analytical Technology - PAT NIRS procedures 14

5. Qualitative procedures **15**

5.1. Development (qualitative procedures) 15
5.2. Data collection (qualitative procedures) 16
5.2.1. Sample collection and population 16
5.2.2. Composition of sample sets 16
5.3. Calibration and optimisation (qualitative procedures) 17
5.4. Validation (qualitative procedures) 18
5.4.1. Specificity 18
5.4.2. Robustness 18

6. Quantitative procedures **19**

6.1. Data collection (quantitative procedures) 19
6.1.1. Sample collection and population 19
6.1.2. Composition of calibration set, calibration test set and validation set of samples 19
6.1.3. Analysis by the reference method 20
6.2. Calibration and optimisation (quantitative procedures) 20
6.3. Calibration model assessment (quantitative procedures) 21
6.4. Validation (quantitative procedures) 21
6.4.1. Specificity 22
6.4.2. Linearity and range 22
6.4.3. Accuracy 22
6.4.4. Precision 22

6.4.5. Robustness 23
6.4.6. Limits of detection and quantification 23
7. NIRS procedure and post-approval requirements 23
7.1. Management of NIRS procedures 23
7.2. Changes to approved NIRS procedures 23
7.2.1. Changes within the defined scope of the NIRS procedures 24
7.2.2. Changes outside of the defined scope of the NIRS 24
7.3. NIRS transfer between NIRS instruments 24
8. Definitions 26

Executive summary

Near Infrared Spectroscopy (NIRS) is an analytical procedure, usually requiring chemometric statistics, with a wide and varied use in pharmaceutical analysis. This includes identification, qualification and assay of starting materials, intermediates and finished products and verification of physicochemical properties.

NIRS also constitutes one of the major techniques in Process Analytical Technology (PAT) and may also be used as part of a Real Time Release Testing (RTRT) strategy.

NIRS requires product understanding. The application of Quality by Design (QbD) principles given in ICH Q8, Q9 and Q10 is considered appropriate with the level of development and/or validation proportionate to the control strategy.

This document provides guidance on the use of NIRS procedures including development, calibration and validation, when used with chemometric statistics, for qualitative and quantitative analysis and in PAT applications.

The development and implementation of an NIRS procedure is iterative and ongoing, and is amenable to the application of lifecycle concepts. These allow good change control practice. Guidance on change control (whether or not within the remit of GMP), taking into account the defined scope of the NIRS procedure, is provided.

A comprehensive 'definitions' section is provided at the end of this guideline.

1. Introduction

Near Infrared Spectroscopy (NIRS) differs from conventional analytical techniques such as HPLC or GC.

For NIRS, the sampling interface and/or probe are essential components. These allow for fast, non-destructive measurement of materials with little or no sample preparation, for a wide range of pharmaceutical forms and manufacturing processes. NIRS enables enhanced assessment of the quality of materials and processes by extensive and more representative sampling.

Interpretation of the complex spectra of unprepared samples generated by NIRS measurement usually requires the use of chemometric calibration models. These models are developed using carefully selected and representative samples, which normally require qualification by independent, reference analytical procedures (normally requiring destructive sample preparation to extract or isolate the analyte of interest and calibration and validation using analytical reference standards).

This guideline aims to relate how NIRS should meet the requirements placed upon all pharmaceutical analytical procedures, taking into account the use of chemometric statistics and that NIRS procedures also need to be robust with respect to the expected variability of the materials or products to be analysed and the manufacturing processes used to prepare them.

The guideline first describes the general requirements for an NIRS procedure in Section 4. This is supplemented by guidance that is specific to qualitative and quantitative procedures in Sections 5 and 6 respectively. Section 7 describes the management of NIRS procedures and post approval requirements.

To aid the narrative of this guidance, the following key terms are used:

NIRS *method:* describes the key elements, principally within the NIRS apparatus, which enable NIRS measurement of the analyte of interest;

NIRS *model:* describes how the NIRS spectral data measured using the NIRS method are related to the analyte of interest, generally employing chemometric software;

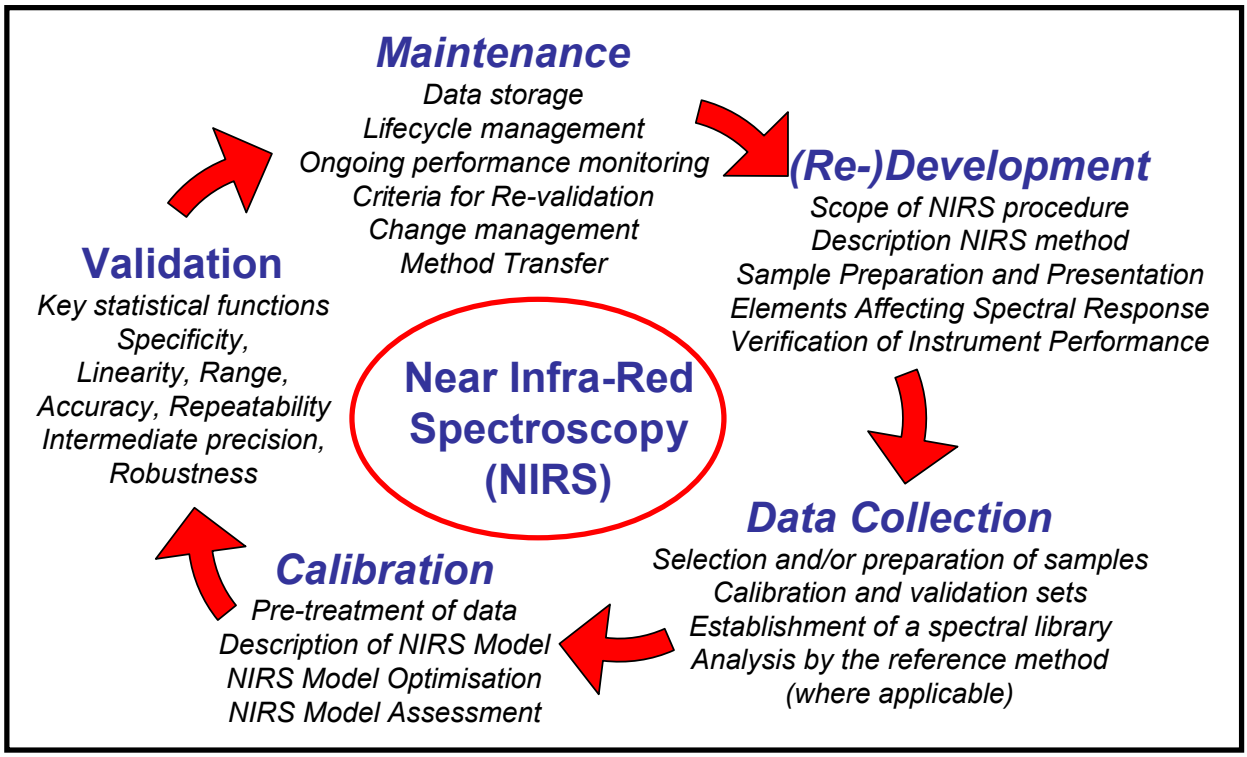
NIRS *procedure:* describes how the NIRS method and model are used for the intended purpose, within its defined *scope.*

It is recognised that the development and implementation of a NIRS procedure is iterative and ongoing, with interdependent component stages, summarised in Figure 1. Some of these stages may not be necessary for all procedures such as variations or PAT applications, for which the absence of information for these stages may then be justified.

The guideline introduces the concept of the **NIRS procedure scope** (Section 4.1.1) to facilitate continuous improvement and life cycle management. Changes *within* the approved scope of the NIRS procedure would be subject to GMP only. Changes *outside* of the approved scope of the NIRS procedure would be subject to variation application (Section 7).

As NIRS procedures cannot be repeated easily by official control laboratories, any reference methods and corresponding specifications should remain in the authorised specifications, with an indication that these methods will not be used for routine batch release.

Figure 1. The iterative nature of NIRS



2. Scope

This guideline describes the regulatory requirements for marketing authorisation applications and variation applications submitted for medicinal products for human or veterinary use, which include the use of NIRS.

NIRS is described in the European Pharmacopoeia; however a single reference to the Ph.Eur. general chapter on NIR spectroscopy (Ph.Eur. 2.2.40) as a sole description for the NIRS procedure is insufficient to support the use of such a procedure in marketing authorisation applications or variation submissions.

This guideline outlines the requirements for applications in which NIRS is used for qualitative and quantitative analysis or in PAT. Approaches other than those described in this guidance may be used, if appropriately explained and justified.

The chemometric principles described within this guideline may also be applicable to other analytical techniques.

NIRS for non-regulatory purposes, such as generating process knowledge, is out of scope of this guideline.

3. Legal basis

This guideline should be read in conjunction with Directive 2001/82/EC and Directive 2001/83/EC.

This guideline should be read in conjunction with:

- Ph. Eur. Monograph 2.2.40 Near-infrared Spectroscopy and 2.9.47 demonstration of uniformity of dosage units using large sample sizes;
- ICH Q2(R1) guideline on validation of analytical procedures (CPMP/ICH/381/95);
- VICH guidelines GL1 & GL2 on validation of analytical procedures (CVMP/VICH/590/98 & CVMP/VICH/591/98);
- CHMP and CVMP guideline on process validation information and data to be provided in the regulatory submission (EMA/CHMP/CVMP/QWP/70278/2012-Rev 1);
- ICH Q8: guideline on pharmaceutical development;
- ICH Q9: guideline on quality risk management;
- ICH Q10: guideline on pharmaceutical quality system;
- ICH guideline Q8, Q9 and Q10 - questions and answers (CHMP/ICH/265145/2009).

4. General requirements

The following is a summary of the general data requirements for NIRS procedures. Additional specific requirements for qualitative and quantitative procedures are given in Sections 5 and 6 respectively.

- the scope of the NIRS procedure (Section 4.1);
- details of the composition of the calibration set, calibration test set and validation set of samples, with justification (Sections 4.2, 5.2.2 and 6.1.2);

- data relating to the reference analytical method, when applicable (Sections 4.2.4 and 6.1.3);
- report(s) of the calibration and validation of the NIRS procedure (Sections 4.4, 4.5, 5.3, 5.4 and 6.2, 6.3, 6.4);
- details as to how the NIRS procedure lifecycle will be managed (Section 7).

If an applicant wishes to use terminology other than that described in this guideline, the terminology should be fully and clearly explained (e.g. using a glossary).

4.1. Development

NIRS has a wide range of qualitative and quantitative applications and its use requires a sound understanding of the physicochemical basis on which its measurements rely and of the instrumental and chemometric principles involved. The applicant should identify any assumptions made during procedure development.

The NIRS signal may be attributed to the analyte of interest or correlated with light scattering effects or with matrix components. The relationship between the NIRS signal and analyte, attribute or process event should be shown to be relevant, scientifically sound and suitable for the intended purpose of NIRS procedure.

The NIRS procedure should be developed to reject samples that are outside of its defined scope (e.g. out of range or compositionally incorrect).

Since NIRS is a rapid, non-destructive analytical procedure, it is possible to enhance assessment of the quality of the materials under test by extensive and more representative sampling. This is consistent with the QbD principle of continuous improvement, and should be taken into account in the development of the NIRS procedure.

4.1.1. NIRS procedure scope

The scope of the NIRS procedure is particularly important to facilitate continuous improvement, and to manage how future changes to the procedure may be implemented from a regulatory perspective (Section 7).

The scope of the NIRS procedure should be clearly identified in any application in which NIRS is used.

The following should be considered in the development of the scope of an NIRS procedure:

- the analyte, attribute or process event(s) to be determined and associated matrix (or matrices);
- the intended purpose of the procedure within the context of the control strategy for the material or product. This should include the limits of the operating range of the NIRS procedure e.g. analyte concentration, manufacturing process parameters and/or design space;
- the NIRS method i.e. the key elements that enable NIRS measurement (Section 4.2.1). This should include the sampling interface, probe position and sampling plans. NIRS procedures are usually sample specific, so sample type and character should be defined (Section 4.2.2);
- the NIRS model i.e. how the NIR spectral data are related to the analyte or property of interest. This should include the relevant statistical acceptance criteria and/or attributes;
- summary information that ensures critical appraisal and revalidation of the NIRS procedure on a regular basis, enabling continuous improvement and appropriate change control, when necessary;

- if used, the reference method should be specified because this is essential to the qualification of samples used for NIRS procedure calibration and validation.

4.1.2. Description of the NIRS method

Summary details of the NIRS apparatus should be provided, including the make and model number of the instrument, the light dispersion principle of the optical system (e.g. grating, FT-IR), the detector type (e.g. silicon, lead sulphide), the measurement method or mode (e.g. reflectance, transmission, transflectance) and the wavelength (wavenumber) range used.

Sample preparation, presentation, sampling devices and details of any other additional components or controls considered necessary for the proposed method should be described.

The number of scans recorded per sample and spectral pre-treatments should be stated and justified.

A spectral library and the means of data collection and storage should be established (see 4.2.5).

4.1.3. Description of the NIRS model

Chemometric data analysis and modelling, using statistical software packages, are usually necessary because of the complex informative content of the NIRS signal. These visualise and extract relevant information from the spectra, or show that the spectra are correlated to a signal measured with the reference method.

Chemometric data analysis works by correlating the variance in the NIRS signal to a number of latent variables or factors, constrained by a set of calibration reference data. There is always a risk that the correlations identified by the software are due to chance only and not to changes in the analyte; therefore chemometric models should always be validated with an independent set of samples.

Details of the NIRS model should be provided, including the commercial software product and the chemometric algorithm (such as PCA or PLS). This should include the relevant statistical acceptance criteria, which should be specified and limited to avoid a drift into unsuitability during the NIRS procedure life cycle. For example, for quantitative methods, the relevant statistical spectral quality test criteria (Section 4.3.1), number of latent variables, Standard Error of Calibration (SEC) and Standard Error of Prediction (SEP, Section 6.3) should normally be stated.

4.1.4. Feasibility study

It is recommended that the feasibility of using NIRS for the intended purpose is considered in development; however, the results of such feasibility studies need not be provided in regulatory submissions.

The feasibility of using NIRS should be considered in the development of new procedures to demonstrate that it is suitable for the intended purpose. Such a feasibility study may include (but is not limited to), the determination of a suitable NIR response, investigations into specificity and matrix interference and the effects of sample handling and preparation.

Other requirements for NIRS procedure development, calibration and validation should also be considered at this stage. For example, for tablet content uniformity testing, the weight of the tablet

should also be considered to ensure measurement of active substance as content rather than concentration.

4.1.5. Risk assessment of variables affecting NIRS procedures

It is appropriate to undertake a comprehensive assessment, according to the principles and means given in ICH Q9, of those risks that may adversely affect the performance of the NIRS procedure in delivering valid results, which may then lead to an incorrect assessment of quality and/or "false positives".

The risk assessment should address the assumptions made in the development of the procedure and its scope, the variables affecting response, and sample presentation.

Each known potential variable that may affect the spectral response should be considered and discussed in turn and either shown to be insignificant or controlled satisfactorily (supported by appropriate data).

It is not possible to list all possible variables, but these may include:

- the environment in which measurement takes place;
- optical quality of the glassware;
- cleanliness of the sample interface;
- optical path length;
- sample optical properties, temperature, thickness;
- sample flow;
- polymorphism;
- particle size;
- residual moisture and solvents;
- homogeneity;
- the age of the samples;
- time of measurement and instrumental drift.

Variability of relevant critical process parameters and critical quality attributes should also be considered e.g. for tablets – granule drying, control of moisture content, compression and tablet hardness.

Other issues that should be considered include spectral (e.g. analyte signal - absolute and relative with respect to matrix and matrix effects) or modelling issues (e.g. co-linearity).

An assessment should be undertaken of the suitability of the number of samples displaying variability and their distribution to the calibration or validation batches. Sampling plans should be representative of the product batches to be marketed and suitable to ensure satisfactory calibration and validation.

A design of experiments (DOE) may be of considerable value as a means to identify high risk elements that require control and risk mitigation.

Risk assessment should be on-going, with critical re-appraisal and revalidation of the procedure on a regular basis. This will support continuous improvement and appropriate change control, when necessary.

A summary of the risk assessment should be submitted to support the development of the NIRS procedure and procedure robustness (Sections 5.4.2 and 6.4.5).

4.2. Data collection

4.2.1. Sample preparation and presentation

Details of sample preparation, if any, should be provided and justified.

Before any NIRS measurement takes place, it is important to optimise the presentation of the sample to the NIRS instrument. Examples of variables that should be optimised are sample orientation, sample size, optical quality of glassware, temperature, material flow, optical path length and environmental conditions.

The interface between the sample and the NIRS detector should be described. The impact of possible variations of the presentation of the sample on the NIRS response should be discussed, supported by appropriate data, and, if shown to be significant, demonstrated to be controlled satisfactorily.

For on-line methods (e.g. determination of homogeneity in blending), it should be demonstrated that the presentation of the sample to the sampling device (e.g. probe) has been optimised and validated. The following should be considered for such applications (the list is not exhaustive):

- optimisation of the sampling device location;
- estimation of the effective sample size;
- assurance that the sampling device interface or window is covered with sample;
- controls to ensure that there is no fouling on the sampling device.

4.2.2. Sample population

Samples should be clearly defined and justified. Sampling procedures should not adversely affect quality of the samples or introduce error or bias.

Samples for calibration and validation should be representative of the commercial production process and the expected variability within the scope of the NIRS procedure. Such variation may include:

- concentration of the analyte of interest;
- particle size;
- material suppliers;
- water content;
- residual solvent content;
- qualitative and/or quantitative variations in the matrix (e.g. excipient grade, formulation);

- process variation (e.g. samples collected over an extended period);
- sample age;
- temperature.

Development and pilot scale batches, which are representative of the commercial process (i.e. same composition, process and quality attributes), may also be utilised.

The validation set of samples should be completely independent of the calibration set (see Sections 5.2.2 and 6.1.2).

Each calibration and validation result should also be independent e.g. for repeatability, each result should be derived independently i.e. from a sample that has been analysed by the NIRS procedure independently from the other samples in the calibration or validation sets.

4.2.3. Pre-treatment of data

Given that NIR spectra are affected by physical parameters such as particle size and sample presentation, raw NIR spectra are often treated mathematically prior to development of the calibration model. Such treatments include normalisation and derivation, which are performed in order to minimise unwanted sources of variation from the data prior to calibration and to enhance spectral features.

Caution must be exercised when performing any pre-treatments because artefacts can be introduced or essential information lost. Any pre-treatment of data should be documented and justified.

A partial wavelength range may be used instead of the complete NIR range. If a partial range is used, it should be justified from a scientific perspective, with data provided to show that all relevant chemical/physical information will be captured.

4.2.4. Analysis by the reference analytical procedure

Samples used for NIRS calibration and validation usually require authentication, or quantitative values to be assigned to them by a reference analytical procedure. When this is not the case this should be explained and justified.

Analysis by the reference procedure and NIRS scanning should be performed on the same samples, within the same appropriate period. More than one reference analytical procedure may be required to authenticate the material under test (e.g. in the case of different grades of materials).

The analyte or attribute determined by the reference and NIRS procedures may be the same e.g. moisture content determination by NIRS and Karl Fischer titration (where data from both methods correspond to water present in the sample) or different e.g. moisture content by NIRS and by percentage loss on drying (where moisture content is assumed to be correlated to the loss in sample mass).

The attributes determined by the reference and NIRS procedures may also differ in terms of units e.g. NIRS analyte as a function of concentration; reference analyte as a function of mass. The additional measurements, e.g. mass or volume of sample, required to convert results of one method to the other should be described.

The use of a reference analytical procedure should be supported by:

- a description of the analytical procedure according to Module 3.2.P.5.2[1] data requirements;
- details of the validation of the analytical procedure according to the Module 3.2.P.5.3[1] data requirements and the ICH Q2(R1) Guideline on Validation of Analytical Procedures (CPMP/ICH/381/95). For Veterinary applications, reference is made to the VICH Guidelines GL1 & GL2 Validation of Analytical Procedures (CVMP/VICH/590/98 & CVMP/VICH/591/98);
- details of relevant reference standards and materials according to the Module 3.2.P.6[1] data requirements;
- The potential for error and bias in measurements between the NIRS and reference procedures should be discussed and addressed, as necessary.

4.2.5. Establishment of a spectral library

The composition of the spectral reference library should cover the scope of the NIRS procedure and should be subject to a change management system (subject to GMP).

Samples should be representative of the marketed materials or products and laid down in a list of batch numbers. The spectra should be indexed such that the source/origin batches can be identified unambiguously.

For qualitative analysis, where the spectral reference library may be very large or diverse, it may be useful to divide the library into appropriate 'sub-libraries' to simplify development. The choice of subsets and the number of sub-libraries should be described and justified.

4.3. Data interpretation

4.3.1. Statistical spectral quality test

Before an NIRS model is applied, each sample spectrum should be subject to a statistical spectral quality test to determine whether the characteristics of the sample fall within the range of variation for which the model was calibrated and validated. In practice, such tests (e.g. hotellings T^2 or distance to model plots) show whether the spectra for the sample fall within a pre-defined range of variation or if the sample is an outlier. The steps taken to address spectral outliers should be described (see Section 4.3.2).

If a sample fails this spectral quality test, it is poor scientific practice to test the sample using the developed NIRS model regardless, since a 'false' positive or otherwise invalid result would be obtained.

The spectral quality test may be performed automatically by the computer software but should nevertheless be understood by the applicant and described, and should include details of the samples chosen to establish the test and an explanation as to its suitability.

[1] Or equivalent in the notice to applicants format for veterinary dossiers.

4.3.2. Outliers in sample data

At any stage of NIRS procedure (re-)development, calibration or validation, systems and procedures should be in place to ensure that the handling of outliers in the data is performed properly. Any suspected outliers in the sample data (NIRS or reference data), which are to be included in the calibration, calibration test or validation data sets should be investigated and any exclusions justified.

The term "NIRS outlier" designates a sample that produces a NIRS response that differs from samples in the calibration model. This does not necessarily indicate an out of specification (OOS) result of the batch. An outlier may be a 'spectral outlier' (i.e. does not comply with the spectral quality test), a 'prediction outlier' (i.e. the spectral data comply with the spectral quality test but the prediction data are out of specification) or the spectral and prediction data are both beyond the proposed scope of the NIRS procedure. An NIRS outlier sample may still meet specifications for the analyte of interest. Such a result may not necessarily be a false observation but merely an observation that is different from the rest and that could have an undefined influence on the results of the analysis.

If a sample is shown to be an outlier because of characteristic properties, the sample should be tested and verified using the reference method and any appropriate alternative analytical procedure. After confirmation of authenticity, the sample may be included in the spectral library and the model should be re-validated so as to include this source of variation. This is an important part of the NIRS procedure lifecycle (see Figure 1, Section 1 and Section 7) and it is important to ensure that the procedure is updated and optimised.

4.3.3. Out Of Specification (OOS) results in routine batch analysis

For routine batch analysis by NIRS, an OOS result should trigger an investigation of the affected batch under the company's pharmaceutical quality system. Rejection or release decisions are based on the outcome of the failure investigation which may include analysis by the reference method.

If, on investigation, the OOS root cause is related to the measurement (human error or instrument failure) or unforeseen spectral variability but the affected batch complies with the specification using the reference analytical method, it can be concluded that the sample meets its specification and the batch can be released. The investigation may also indicate that the NIRS procedure has not been fully developed. The NIRS procedure should be re-evaluated and updated to include the new source of variability, as appropriate (as per the NIRS procedure lifecycle concept, see Figure 1).

While the NIRS procedure is being updated, its use should be suspended: the commercial manufacture can proceed using the reference or other registered analytical procedure.

Repeated changes from NIRS to reference method should be avoided and an OOS result by the NIRS procedure should always lead to an investigation rather than an immediate referral to the reference method.

4.4. Calibration

For NIRS, calibration is generally performed by developing a calibration model that relates concentrations or properties to absorbance spectra for a set of reference samples (the reference library or the calibration set). Specific requirements for calibration are described in the sub-chapters for 'qualitative procedures' and 'quantitative procedures'.

4.5. Validation

The validation data is to be used to assess the suitability of the calibration model and to comply with the data requirements for Module 3.2.P.5.3[2] and the guidance given in ICH Q2(R1) guideline on validation of analytical procedures (CPMP/ICH/381/95) and, for veterinary applications, GL1 & GL2 validation of analytical procedures (CVMP/VICH/590/98 & CVMP/VICH/591/98).

The acceptance criteria for calibration model assessment and validation should be specified and justified with reference to the intended purpose of the NIRS procedure.

A comparison of results obtained by analysis of the same set of samples by the NIRS procedure and the reference analytical procedure (if applicable) forms part of the validation of NIRS, along with independently determined parameters, such as intermediate precision.

The uncertainty of measurement, e.g. accuracy and precision, of the reference analytical procedure should be taken into account when generating validation data for the NIRS procedure.

If the NIRS procedure is being presented in the initial registration dossier, validation data should also be presented for the reference analytical procedure, if used (see Section 4.2.4).

If the NIRS procedure is being registered as a variation to a marketing authorisation for which a reference method is already approved, then a summary of the validation data for the reference method, in compliance with the current (V)ICH guidance on validation of analytical procedures, should be provided.

Specific requirements for validation are described in the sub-chapters for 'qualitative procedures' and 'quantitative procedures'.

4.6. Process Analytical Technology - PAT NIRS procedures

Because PAT NIRS procedures are specific to the nature of the manufacturing processes (e.g. sampling frequency adapted to process dynamics), it is not appropriate to prescribe exact requirements for such procedures in this guideline.

The general data requirements described in this guideline are also applicable to NIRS PAT procedures and the amount of information required for an NIRS PAT procedure will depend on its intended purpose and scope. The criteria necessary to ensure that the NIRS procedure is valid and fit for purpose should be justified.

The PAT NIRS procedure should be described in detail, with specific information in relation to the applied PAT application (e.g. spatial placement of the devices, structural peculiarities of the manufacturing facility and the nature and extent of sampling).

An example of the use of NIRS in a PAT application is the monitoring of a powder blend for homogeneity. The blend may be monitored in terms of the measurement of the change of the NIR signal (e.g. its standard deviation) over time (also called moving block standard deviation (MBSD)), where this has been shown to be a valid measure of homogeneity.

[2] Or equivalent in the notice to applicants format for veterinary dossiers.

5. Qualitative procedures

5.1. Development (qualitative procedures)

NIRS has a wide range of qualitative applications, almost all of which could be divided into the three areas of identification, qualification and conformity checks.

Identification and qualification

This guideline uses the term **identification,** when this is to confirm identity only of a substance, and the term **qualification,** when this is to confirm identity and to differentiate between grades of the same substance (e.g. different grades in terms of particle size or different polymorphs).

If identification and/or qualification are based on more than one analytical method, then it should be clear which reference method(s) will be replaced by the proposed NIRS procedure.

The identification or qualification of a substance (e.g. drug substance, excipient, blend, drug product or intermediate) using NIRS is based on the comparison of the spectral data of the substance with the spectral data of several samples of several batches of different substances present in a spectral reference library. It may be necessary to apply chemometrics to compare the data and set acceptance criteria for a positive/negative (match/no match) identification or qualification. The criteria should be justified.

The identification or qualification of a substance can be performed in several stages. For example, the identification of a chemical identity or a group of related substances may be performed initially, followed by the use of more selective libraries for each individual grade or substance. This approach can be used to decrease the likelihood of false positives/negatives. Qualification is often performed after the sample identification, with the qualification spectral library derived from samples chosen to represent the defined variability of the different grades of the same substance.

When the NIRS procedure on its own is not sufficient to identify or qualify a substance, the scope of the NIRS procedure should be defined as being only one component of a battery of tests required to fulfil identification or qualification, i.e. it should be supplemented by other different analytical procedure(s), so that the tests taken together ensure specificity.

Conformity checks (e.g. PAT, dynamic process monitoring, trend analysis or in-process controls)

This guideline uses the term **conformity** as the conformation of characteristics in accordance with a certain degree of similarity (chemical and/or physical attributes) to a specified standard or event. An example of such a check is the determination of the endpoint of a process by monitoring the change in NIRS signal.

Conformity NIRS procedures may also be known as 'dynamic process monitoring', trend analysis or PAT procedures (see Section 4.6) and may not involve the use of a reference analytical method because of difficulties in sampling for reference analysis.

Conformity checks should be treated in a similar way to qualitative procedures with respect to calibration and validation; however, the extent of the calibration and validation work performed will depend on the intended purpose of the procedure. For example, validation of a moving block standard deviation (MBSD) procedure will focus mainly on a specific end-point, supported by a sound rationale and analytical evidence of the procedure's predictive ability (i.e. the link between end point and true blend uniformity having been established during method validation).

5.2. Data collection (qualitative procedures)

5.2.1. Sample collection and population

The selection of samples and, where necessary, the extent of spectral library development, will depend on the complexity of the procedure. All samples should be verified with the approved reference methods where applicable, or authenticated by appropriate means (certificate of analysis or relevant testing).

Where laboratory or pilot scale samples are required to present wider variability than that shown for production samples, such samples should be prepared using the same manufacturing procedure as used for routine batches, unless otherwise justified. The balance of production to development batches in all sample sets should be justified with respect to the variation expected in routine production. The choice of samples should be sufficient to ensure the robustness of the NIRS procedure for routine use.

For procedures used to identify or qualify substances on receipt, samples from all relevant suppliers used should be incorporated into the library.

For conformity checks, the sample population should be justified with respect to the intended purpose of the procedure.

5.2.2. Composition of sample sets

In order to develop, optimise and validate a calibration model for a typical qualitative NIRS procedure used for identification or qualification, three sets of samples are required:

- a calibration set for creating the calibration model;
- a calibration test set for internal validation and optimisation of the model (cross validation techniques may be applied instead, using the calibration set of samples);
- an independent validation set for external validation of the model.

The number of samples per batch and the number of batches used for calibration and validation should be justified.

Each set of samples should be representative of the intended scope of the NIRS procedure and include samples covering the full range of potential variation in the sample population. The intensity of the analyte signal and the complexity of the sample matrix and/or interference by the matrix of the analyte signal of interest should also be taken into account. In general, the more complex and the more interference from the matrix, the more samples will be required.

The calibration set contains all those samples proposed for inclusion into the spectral library. In the simplest form of spectral library, samples of all materials used at a particular site are included in one library and chemometric analysis is applied to this library. Since this may not always provide adequate specificity, sub-libraries are often used, containing all samples of a particular class, to ensure the required specificity. Identification and qualification may therefore be an iterative process, with identification of a substance in the first instance using the main spectral library followed by qualification of, for example, the polymorphic form of that substance using a sub-library.

The selection of an appropriate calibration model may be aided by so-called 'internal validation' methods. 'Internal validation' is the application of resampling statistics to cross-validate and provide an 'internal check' of the performance of the model for the purposes of optimisation. A subset or

subsets of the spectral reference library data are subjected to a variety of statistical processes to identify which calibration model (generated by the software) may best fit the available data.

A calibration test set may be used (instead of cross validation within the calibration set) to provide the first 'test' or check of the validity of the model. The calibration test set does not represent independent validation of the NIRS procedure (which must be carried out using an entirely independent set of samples), since the samples are taken from different batches within the same (historical) population of batches. In practice, the calibration set often consists of two thirds of the available sample population and the calibration test set is the remaining third, however this may not always be the case and the calibration set should always contain a sufficient number of samples to ensure that the generated calibration model is robust. The applicant should give the rationale for the composition and number of samples in the calibration and calibration test sample sets and justify their suitability.

The external validation set of samples should be entirely independent of those samples used to build the spectral library and should include qualitatively positive and negative samples. This set should not be taken from the same (historical) population as those batches used to generate the calibration model.

In principle, the external validation set should cover the calibration range of the NIRS model, including all variation seen in the commercial process and should include pilot and production-scale batches, where possible.

For conformity procedures, the external validation samples would be expected to cover the range of variability shown within the process being monitored and to give correct determination of an end point. Positive and negative results should be included in the independent validation set of samples, to ensure that the NIRS procedure is fit for purpose (e.g. for assessment of homogeneity in a blending process, the inclusion of samples that have been under-blended or over-blended to the point of de-mixing).

5.3. Calibration and optimisation (qualitative procedures)

Some examples of calibration algorithms are Principal Component Analysis (PCA), discriminant analysis (linear or quadratic), Soft Independent Modelling of Class Analogues (SIMCA), cluster analysis and correlation algorithms such as distance-match or shewart charts.

The selection of the most appropriate algorithm for calibration depends on the scope of the NIRS procedure. In general, the simplest available algorithm that gives successful results should be used.

It is almost always necessary to determine acceptance criteria (e.g. thresholds, confidence limits or tolerances) for identification and/or qualification. The acceptance criteria should be justified. Graphical representation could be useful to support the information provided (e.g. loading plots).

In general, the optimisation of a qualitative procedure is confined to the selection of the samples included in the model, the choice of pre-treatments and the choice of calibration algorithm. Internal validation is often used to 'test' and allow further optimisation. This internal validation step should demonstrate that all samples of the spectral reference library are identified or qualified according to the scope of the procedure, using the defined acceptance criteria.

5.4. Validation (qualitative procedures)

The external validation step will provide evidence that the model can be used in routine analysis for new samples, within the NIRS procedure scope.

The validation of the NIRS model should demonstrate that spectra of an acceptable minimum number of batches have been included in the spectral library and that these batches are sufficiently representative to cover the normal variation of the substance.

The composition of the independent (external) validation set of samples should be described and justified.

The applicant should demonstrate that the NIRS model is suitable for the intended purpose by the means of appropriately defined and justified acceptance criteria.

Alternative thresholds and/or statistical parameters may be used to evaluate the performance of the model. These should be stated, fully explained and their suitability for the intended purpose should be justified.

Qualitative NIRS procedures should be validated for a minimum of specificity and robustness.

5.4.1. Specificity

The extent of specificity testing depends on the scope of the NIRS procedure. A lack of specificity may be compensated for by other supporting analytical procedures (see Section 5.1).

Independent samples of substances represented in the spectral reference library, but not used to create it (e.g. different batches, blends), must be tested. All of these samples should be approved correctly (pass or match), whereas potential challenges should be rejected (fail or no match).

For the identification or qualification of pharmaceutical substances, relevant existing name and structure analogues (if available) should be included in the validation set, unless their absence is justified. A review of the goods-in and manufacturing operations should be used to justify the analogues and challenges presented to the model.

Where applicable (e.g. for qualification procedures), validation for specificity should include challenge with different grades of the same substance, anhydrous/hydrated material, different polymorphs, different particle size or material supplied by different vendors.

The results should demonstrate that for each tested parameter, the NIRS procedure is sufficiently selective to discriminate between batches that comply with the tested parameter and batches that do not, as effectively as the reference method.

5.4.2. Robustness

The effects of relevant variables e.g. temperature (environment and sample), humidity, different position of the sample in the optical window, different sample presentation devices, variation in sample bottles/vials, probe depth or, if applicable, different packaging materials, should be understood, tested and documented. Instrumental variations may also be considered in the validation for robustness, e.g. changing lamps or reflectance standard. The use of Design Of Experiments (DOE) may be considered to maximise the information available.

6. Quantitative procedures

6.1. Data collection (quantitative procedures)

6.1.1. Sample collection and population

Where feasible, samples of production batches should be augmented with those from development batches, manufactured specifically to simulate the limits of potential variation in the sample.

Sub-batches for which manufacturing process parameters are systematically changed during a production run could also be considered to simulate such variation.

Where laboratory samples are required to expand the narrow range of production samples to properly assess linearity in line with specification limits, such samples should be prepared using the same manufacturing procedure, or any changes should be justified with regard to their impact on the NIRS model.

The balance of production to development batches in the sample set should be justified with respect to the variation expected in routine production.

When feasible, a uniform distribution of samples throughout the range of potential variation should be ensured. The distribution of samples should be evaluated with respect to the intended purpose of the procedure. For cases in which the NIRS procedure is used only as a limit test, e.g. water content in a lyophilised powder for injection, it may be acceptable to include more samples around the proposed specification limit. This should be explained and justified.

The possibility of the introduction of undesirable correlations and systematic errors should be considered, taking account of the known NIRS signal of the analyte of interest, spectral selectivity and potential matrix effects.

The choice and number of samples should be justified with respect to the intended purpose of the procedure. Sample collection and population may be addressed effectively and efficiently using an appropriate DOE approach.

6.1.2. Composition of calibration set, calibration test set and validation set of samples

To develop, optimise and validate the calibration model for a quantitative NIRS procedure, the following sets of samples are required:

- a calibration set for creating the calibration model;
- a calibration test set for internal validation and optimisation of the model (cross-validation techniques may be applied instead, using the calibration set of samples);
- an independent validation set for external validation of the model.

Calibration algorithms are generally based on the correlation of variance in NIRS spectra, via their latent variables (or equivalent; see 'definitions' Section), with a reference data set. The number of samples used to develop the calibration model should be much greater than the number of latent

variables used and the number of latent variables should be comparable with the number of detectable significant sources of variability found during the risk assessment step.

The number of samples to be included to create a valid calibration model for quantitative analysis will also depend on the complexity of the sample matrix and/or interference by the matrix of the analyte signal of interest. In general, the more complex the sample matrix, the more samples will be required.

See also Section 5.2.2, as this guidance for qualitative procedures is also applicable to quantitative procedures.

The applicant should give the rationale for the composition and number of samples in each set used for calibration and validation and justify their suitability for establishing a NIRS model that is representative of expected variability of future batches.

An applicant may decide not to make an internal validation assessment during development and optimisation of a calibration model. In such cases, a calibration set and an independent validation set of samples would be the minimum requirement.

6.1.3. Analysis by the reference method

The performance of a quantitative NIRS procedure is dependent on the performance of the reference method. Poor precision and accuracy of the reference method will limit the performance of the NIRS procedure. It is important that care is taken to ensure that uncertainty in the reference method is low in relation to the intended performance of the procedure.

Repeated sample analysis by the reference method should be discussed and reference data should be tabulated and presented graphically. The number of replicates to be averaged to provide reference data for the calibration model should be stated and justified with reference to the performance (precision and accuracy) of the reference method and the NIRS procedure.

Where the reference analytical method determines the amount of analyte in a sample as a different function or unit from the NIR procedure, e.g. mass versus concentration, then result assignment will also require conversion using additional measurements and this should be fully explained.

6.2. Calibration and optimisation (quantitative procedures)

Following acquisition of matched spectral and reference analytical data for the calibration set of samples, the chemometric calibration model should be developed using a specified software package. Such software empirically correlates variation within the data. The result of this correlation may be presented in different ways depending on the algorithm used (e.g. latent variables for PLS regression or principal components for PCR).

The number of latent variables to be used in the calibration model is of critical importance to avoid under or over fitting of the data and to ensure that the model is optimised for the intended purpose of the procedure.

The following should be considered when choosing the number of latent variables to use:

- co-linearity;
- minimal contribution to the data variance arising from spectral variations of the analyte of interest;
- contribution to the data variance arising not from the spectral variations of the analyte of interest, but from other components of the sample matrix, e.g. excipients or other characteristics.

The above list is not exhaustive and relevant issues revealed by the risk assessment and feasibility studies should be taken into account.

Loadings plots, describing the variation explained by each latent variable may be useful when choosing the number of latent variables to use.

Optimisation of calibration models may be performed by 'internal validation' methods as described in Section 5.2.2.

The use of resampling statistics for optimisation is a rapidly developing field and more appropriate statistical processes may be available to assess under- and over-fitting of models. Any process used should be explained and justified.

6.3. Calibration model assessment (quantitative procedures)

To support a claim that an appropriate number of latent variables have been used in a calibration model, the Standard Error of Calibration (SEC) and the Standard Error of Prediction (SEP) should be plotted as a function of the number of latent variables. The proposed calibration model should be characterised by graphical plots of reference values against NIRS predicted values of *both* the calibration and validation sets of samples, to give a visual overview of linearity, bias, slope and outliers for *both* sample sets. Plots of residuals against NIRS predicted values should also be provided, to provide an additional visual assessment of linearity.

For inverse multivariate regressions (e.g. PLS), the reference values should be plotted as the dependent variable and the predicted NIRS values plotted as the independent variable.

For both sets (calibration and validation), outliers should be assessed as explained in Section 4.3.2 and relevant statistical parameters should be reported. When an outlier is justifiably removed from the set of samples, then the calibration and validation data should be reprocessed.

It is expected that a good correlation coefficient is obtained (close to 1), with slope, bias and intercept not statistically different from 1, 0 and 0 respectively.

For the calibration set, the SEC or equivalent should be reported.

For internal (cross) validation methods of optimisation, the Standard Error of Cross Validation (SECV) should be reported. Other statistical parameters may also be used. Any statistical parameter(s) used should be stated, explained and justified.

For the validation set, the SEP should be reported. The calibration range should be at least 10 x SEP. The mathematical definition of SEP as defined in Section 8 of this guideline does not contain a bias correction, in line with expectations for pharmaceutical applications. The definition of statistical parameters used by applicants should be stated in the dossier.

The calibration set and validation set of data should be compared and differences discussed in terms of the suitability of the calibration model. It may be appropriate to report and justify the SEP/Standard Error of Laboratory (SEL).

6.4. Validation (quantitative procedures)

The NIRS model will be valid only to cover the variability, e.g. range, used in the validation of the model.

The NIRS procedure should be able to reject samples that are outside of its defined scope, such as out of specification product, placebo, samples containing different quantitative composition of proposed excipients, and samples containing different active substance and excipients (see also Section 4.3.1; spectral quality test).

The accuracy and precision of the reference analytical procedure should be taken into account when generating validation data for the NIRS procedure.

6.4.1. Specificity

The NIRS procedure should be able to assess unequivocally the analyte in the presence of other components, which may be present.

The following may be used as supportive evidence of specificity:

- reference to the development and risk assessment data demonstrating a suitable NIR response based on the known NIR characteristics of the analyte;
- comparison of the loadings plots for the components used to develop the chemometric model, against the known NIR characteristics of the analyte;
- statistical spectral quality test (Section 4.3.1);
- demonstration that samples that are outside of the NIRS procedure scope are rejected;
- validation data to demonstrate accuracy and robustness.

6.4.2. Linearity and range

To demonstrate linearity, it is required that samples in the validation set are distributed across the specified range.

The NIRS results should be compared with those of the reference method (if applicable). The correlation coefficient and analysis of residuals (indicators of linearity), should be explained, justified and supported by graphical representation.

The applicant should justify the choice of statistics applied to determine linearity, if these differ from those described in this guideline.

6.4.3. Accuracy

Accuracy should be established across the specified range of the NIRS procedure, which would normally be by comparison of results with the validated reference method. Bias should not be statistically different from zero.

6.4.4. Precision

Repeatability and intermediate precision should be determined, covering the specified range.

The suitability of the determined precision of the NIRS procedure should be discussed and justified.

6.4.5. Robustness

Evidence to demonstrate the robustness of an NIRS procedure should cover chemical and physical sample variables, the conditions employed for sampling and sample preparation, as well as variations in procedure parameters.

If these have been considered, reference to data generated from the development and optimisation of the calibration model and the validation data described above would be supportive. Otherwise, validation data for the determination and assurance of robustness should be provided.

6.4.6. Limits of detection and quantification

Limits of detection and quantification for the proposed NIRS procedure need only to be demonstrated when relevant and where the analyte is considered to be an impurity (e.g. water content).

7. NIRS procedure and post-approval requirements

7.1. Management of NIRS procedures

It is recognised that NIRS procedures will evolve over time (see Figure 1, Section 1). The applicant should indicate how they will manage the NIRS procedure lifecycle in the initial application, since this is essential for the efficient and effective implementation of post-approval changes to the NIRS procedure.

NIRS procedure lifecycle management should ensure critical appraisal and revalidation of the NIRS procedure on a regular basis, and enable continuous improvement and appropriate change control when necessary during its life cycle. This should be addressed in the development of the scope of the NIRS procedure (see Section 4.1.1).

It is possible to update calibration models of NIRS procedures as new data become available following the purchase or production of new analyte batches. This is considered a good means of continuous improvement and is recommended.

Any extension of the scope of the NIRS procedure should be implemented by variation application only.

Changes with respect to the product lifecycle, e.g. changes in manufacturing processes, composition, site of manufacture, would normally also require that the NIRS procedure is updated.

7.2. Changes to approved NIRS procedures

Changes (both planned and unplanned), which might affect the performance of an NIRS procedure, may necessitate re-calibration and/or re-validation of the NIRS model to demonstrate continued model suitability. All changes should be validated accordingly, appropriately documented and recorded according to valid change management protocols.

7.2.1. Changes within the defined scope of the NIRS procedures

In general, changes *within* the scope of the NIRS procedure would be subject to GMP only. Relevant examples include the maintenance of the spectral library and replacement of equipment consumables with similar, including lamps, sampling devices, location and software upgrades.

Changes should be fully documented, and should include appropriate re-validation and comparability reports to show that the revised NIRS procedure is consistent with the approved procedure. A risk assessment should be conducted to determine the risk associated with the change being made.

For qualitative NIRS procedures, suitable change management tests (tests used to demonstrate unchanged NIRS procedure reliability following a change) should be in place for each NIRS procedure and spectral reference library (where applicable). A change management test should be composed of a minimum of two sets of samples (i.e. two classes or substances) for which separation is most critical. If the NIRS procedure does not comply with the change management test (meaning that the procedure is unable to distinguish between the two sets of samples), the model should be fully re-validated. The suitability of the change management test should be subject to periodic re-evaluation.

Quantitative NIRS procedures should only be used within the calibrated concentration range and using conditions defined in calibration. It may be appropriate to add sample observations into the calibration model (within the calibrated range detailed in the defined scope of the NIRS procedure) via model updates. Such changes require re-validation and documentation should be available for GMP inspection.

7.2.2. Changes outside of the defined scope of the NIRS

Extensions *outside* of the approved scope of the NIRS procedure are subject to variation application, which should include an appropriate comparison of the updated NIRS procedure with the current procedure and/or reference analytical method, if applicable.

For extensions of the scope of a *qualitative* NIRS procedure e.g. to apply a previously approved *qualitative* NIRS procedure to other authorised drug substances and/or drug products, a statement of compliance with this guideline and a comparison of the updated NIRS procedure with the current procedure would be considered sufficient.

For *quantitative* analysis, extensions of the scope of NIRS procedures include for example, changes of ranges and/or specification limits. Variation applications for such changes require evidence of re-calibration and validation of the NIRS model.

7.3. NIRS transfer between NIRS instruments

The aim of NIRS procedure transfer is to ensure that the calibration model generated on one NIRS instrument will work on another instrument, based on the validation parameters detailed in Sections 5.4 and 6.4 of this guideline. The following parameters are essential to judge the similarity of instruments:

- hardware (e.g. identical spectrometer type and measuring set-up);
- software (including mathematical algorithm and how the spectra are treated in the calibration model);

- interfaces (e.g. probes, waveguides and fibre optics);

Depending on the scope of the NIRS procedure (e.g. qualitative, trend analysis or quantitative) and the degree of similarity of the instruments involved, NIRS procedure transfer will be performed using option 1 or option 2 below:

1. In cases for which there are NIRS method changes only:

- It may be possible and sufficient to include a mathematical compensation in the applied software of the NIR instrument to ensure that identical spectral responses are achieved for a representative set of reference samples when tested by all instruments. In general, a bias and/or slope or vector correction is undertaken.

2. In all other cases, for which mathematical compensation is insufficient:

- Calibration and validation of the NIRS model should be repeated and confirmed on the additional instrument(s).

- Calibration transfer models may be developed using a small but representative number of calibration samples that are run on both instruments (the master and the additional instrument). A convenient method for choosing samples is one that is based upon their good multivariate leverage. In this method, samples are selected that have a large influence on the calibration model. Depending on the complexity of the multivariate model, a smaller representative number of samples (in comparison with the number used to calibrate and validate the model originally) should be sufficient to support model transfer between instruments.

- In the event that the master instrument is no longer available, an appropriate number of samples should be justified and used to build the calibration model on the additional instrument(s).

In both cases (1) and (2), the transfer of an NIRS procedure to another instrument should be the subject of an appropriate comparability protocol to demonstrate successful transfer.

The comparability protocol for NIRS procedure transfer(s) may be submitted with the initial application as a 'post approval change management protocol' if changes are foreseen in the applicant's strategy. Alternatively, the comparability protocol may be submitted at the time of the variation application to register the transfer itself. The protocol should include criteria that have been justified to be acceptable to demonstrate a satisfactory transfer.

8. Definitions

Bias (mean of the errors)	A statistic measuring the mean of the errors between the NIRS and reference quantitative analyte values	
	$Bias = \frac{\sum_{i=1}^{n}(y_i - Y_i)}{n}$	*Y* = NIRS predicted value *y* = reference method value n = number of samples
Calibration set	The set of samples used for creating the calibration model	
Calibration test set	The set of samples, which are drawn from the same population as the calibration set, but were not used to generate the calibration model. In practice, the calibration set often consists of two thirds of the available sample population. The calibration test set is the remaining third	
Calibration test set (internal validation)	The application of possible chemometric calibration models to the calibration test set to perform an internal validation or 'check' of the calibration model.	
Change management protocol	A protocol listing potential future changes in the NIRS procedure and the actions considered necessary to prove the maintained reliability of the procedure (ref. EMA/CHMP/CVMP/QWP/586330/210 Q&A on post approval change management protocols)	
Chemometrics	Mathematical methods to analyse and compare data	
Cross-Validation	See Internal validation	
Conformity	Characteristics in accordance with a certain degree of similarity (chemical and/or physical entities) to some specified standard	
Design of Experiments (DOE) (factorial experimental design)	A structured, organised method for determining the relationship between factors affecting a process and the output of that process (also known as 'formal experimental design')	
Identification	Determination of a chemical identity	
Internal validation	Subsets of the calibration data set are subjected to a variety of statistical processes to identify which calibration model may best fit the available data. Each model is characterised by a statistical parameter. For cross-validation, the entire data set of samples is split into individual samples or groups of samples, which are removed individually from the rest of the samples and tested as unknowns against a calibration model constructed using the rest of the samples. The characteristic statistic is the Standard Error of Cross Validation (SECV)	
Latent variable	Latent Variables are calculated by a mathematical model (e. g. PCA) on the basis of values of observed variables with the aim of covering the maximum variance of the observed data and reduction of dimensionality and/or covariance of the observed data (e. g. spectral data). Examples for latent variables are 'factor' or 'principal component' depending on the mathematical method which is used.	

Leverage	In chemometrics the leverage is a concept related to the Mahalanobis distance and is used to measure the influence of a sample in a model based on its similarity to the rest of the population. The Mahalanobis distance takes into account the correlations of the data set and is scale-invariant, i.e. not dependent on the scale of measurements. The leverage of a sample is the distance to the centre of all samples relative to the variability in its particular direction
Loading plot	The loading plot for each principal component indicates the magnitude (small or large correlation) and the manner (positive or negative correlation) of how each original measured variable (e.g. wavelength of the NIR spectra) contributes to the variance seen for the analyte signal
Multiple Linear Regression (MLR)	A mathematical method that may be used in the quantitative analysis of NIR spectra, in which the absorbance (A) of samples at one or more wavelengths is correlated to reference values.
NIRS Procedure	Describes how the NIRS method and model are being used for the intended purpose.
NIRS Method	Describes the key elements that enable the NIRS measurement of the analyte of interest. This includes for example, the equipment and spectrophotometer type (e.g. FT, grating etc), the sample measurement interface (e.g. probe, sample stage etc), the number of scans or measurements and the spectral range of the instrument.
NIRS Model	Describes how the NIR spectral data are related to the analyte property of interest or the intended use of the procedure.
No match conclusion	The sample is not considered identical to any entity in the reference library
Pass conclusion	The sample is considered identical to an entity in the reference library
PCA	Principal Component Analysis
PCR	Principal Component Regression
PLS (PLSR)	'Partial Least Squares' (Regression).
Pre-treatment	Processing of the spectral data, with mathematical or other techniques, prior to chemometric analysis
Principal component	Principal components are calculated by means of chemometric software from a set of original variables (e.g. NIR spectra) by linear transformation of the original variables into a lower dimensional space. The principal components have the characteristic that a maximal amount of information about the original variables is retained
Process Analytical Technology (PAT)	A system for designing, analysing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality
Qualification	1. Characterisation based upon chemical and physical attributes. 2. Determination of the chemical identity and the variability of the sample within the defined variability of the material
Qualitative procedure	Procedure with a yes or no result, e.g. identity
Quantitative procedure	Procedure with a numerical result, e.g. assay

Reference library (spectral reference library)	Database containing spectra of several batches of several substances to be tested. Spectra of unknown samples are compared with this database	
Reference method	The conventional analytical method that is used to determine the concentration or property value of the samples	
Re-sampling Statistics	Statistical methods to aid the optimisation of the calibration model by using subsets of the calibration set, e.g. cross-validation	
Standard Error of Calibration (SEC)	A measure of the variability of the difference between the predicted and reference values for a set of calibration samples, where:	
	$SEC = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{C,i} - Y_{C,i}\right)^2}{n-p}}$	Y_C = NIRS predicted value of calibration set y_C = reference method value of calibration set n = number of samples p = number of coefficients, e.g. wavelength (MLR), principal components (PCR), factors (PLS)
Standard Error of Cross-Validation (SECV)	A statistic measuring the difference between the NIRS procedure and reference method quantitative analyte values of the calibration set using a cross-validation method.	
	$SECV = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{CV,i} - Y_{CV,i}\right)^2}{n}}$	Y_{CV} = NIRS predicted value y_{CV} = reference method value n = number of samples
Standard Error of Laboratory (SEL)	The SEL concerns to the intermediate precision (intra-lab) or reproducibility (inter-lab), whichever is applicable	
	$SEL = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{1,i} - y_{2,i}\right)^2}{n}}$	$y_{1/2}$ = reference method value measured at different laboratory conditions n = number of samples
Standard Error of Prediction (SEP)	A measure of the variability of the difference between the predicted and reference values for a set of independent validation samples, where:	
	$SEP = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{V,i} - Y_{V,i}\right)^2}{n}}$	Y_v = NIRS predicted value y_V = reference method value n = number of samples
Validation set	Independent set of samples used in validating the NIRS model	