TECHNICAL SPECIFICATION

ISO/TS 22176

First edition
2020-01

Cosmetics — Analytical methods — Development of a global approach for validation of quantitative analytical methods

Cosmétiques — Méthodes analytiques — Développement d'une approche globale pour la validation des méthodes analytiques quantitatives

Reference number
ISO/TS 22176:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 217, *Cosmetics*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The purpose of this document is to propose a characterization protocol for the validation of a quantitative analysis method in the cosmetic field and thus responds to the requirements of ISO/IEC 17025, i.e. using the performance goals as a basis. The theoretical principles of this approach can be found in Reference [1]. This document is based on the French Standard NF V 03-110[2].

Analytical methods for analyses of cosmetics need to be validated. Validation has been long considered as a process consisting in individually verifying several different criteria, i.e. selectivity, repeatability, linearity, trueness, etc. The global approach, as proposed since 2003[1], is based on the total error concept and the term "global" means that only a single criterion should be checked to validate a method: the agreement between a future experimental result and the true value. This approach has already been applied in the domains of pharmacy[1],[9], agricultural chemistry[2], and is in agreement with quality assurance guidelines such as GLP or ISO/IEC 17025. This validation process applies generally to already developed methods and includes evaluations of the following criteria: specificity/selectivity, precision, trueness, linearity range, LOD/LOQ, stability, ruggedness.

The large number of cosmetic products and the variety of matrices present a challenge for an analytical laboratory requiring that standardized methods to be adapted for each type of samples. Additional difficulties are linked to the very low concentrations to be measured, generally of the order of the mg/kg (ppm) or µg/kg (ppb). In such context, criteria such as accuracy and uncertainty of measurement of the analytical results are of utmost importance.

When the concentration of a substance is determined by an analytical laboratory, it is important to evaluate the gap between the measured value and the known true value. This difference indicates the trueness of the analysis. If cosmetic samples are analysed several times in different conditions (laboratory, instrument, operator), the individual results will present a dispersal around the average value which represents the precision of the measurement. As for the individual measurement, it represents an error with the average value and an inaccuracy with regard to the reference value (i.e. the true value).

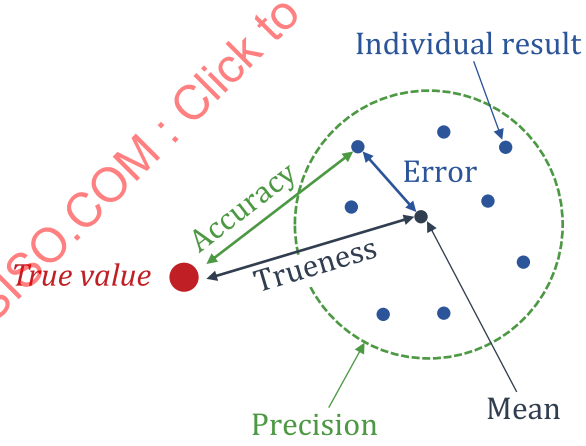
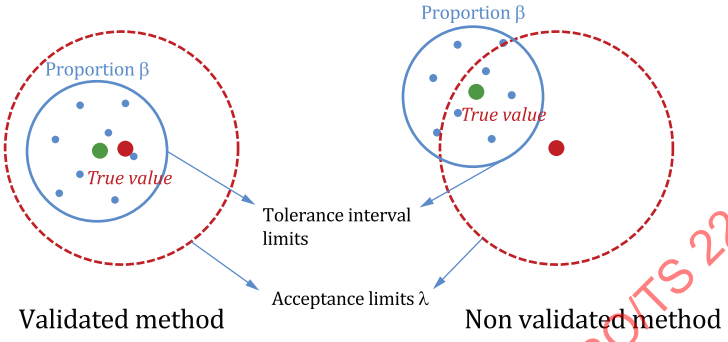


Figure 1 — Illustration of the concepts of accuracy, precision and trueness

When a laboratory measures the concentration of a given substance in a cosmetic product sample, the value which is obtained is thus characterized by a given accuracy which includes at the same time the notion of trueness and precision (see Figure 1). It can also be considered as total error. The insurance that the accuracy of a result is below acceptable limits, is thus one of the ways to make sure of the validity of a measurement.

The accuracy profile (plot of accuracy versus concentration), such as it is developed in numerous domains[3] to [9], is thus the way to know the accuracy on a result obtained with a given method applied to a type of sample in the environment of a given laboratory.

To reach this accuracy profile, it is necessary to undergo a specific assay allowing to demonstrate the validity of the analytical method, as well as the accuracy of the measurement for a given substance. In this approach, it is necessary to determine a tolerance interval[10] which contains a given proportion (β) of future measured values inside (in average). If this tolerance interval is located inside a limit of acceptability defined a priori, taking into consideration several parameters such as the type and concentration of analyte, type of matrix, of analysis and conditions of the experiments, in this case, the method will be considered as valid, and if it goes outside this limit of acceptability, the method will be considered as non-valid (see Figure 2).



Key

- mean value
- true value

Figure 2 — Illustration of the validation principle

Cosmetics — Analytical methods — Development of a global approach for validation of quantitative analytical methods

1 Scope

This document defines a global approach for the validation of a quantitative analytical method, based on the construction and interpretation of an accuracy profile, and specifies its characterization procedure.

This procedure is particularly applicable for internal validation in a cosmetic testing laboratory, but its scope can be extended to the interpretation of data collected for an interlaboratory study designed according to the recommendations of the ISO 5725-1. It does not apply to microbiological trials. The present approach is particularly suited to handle the wide diversity of matrices in cosmetics. This document only applies to already fully-developed and finalized methods for which selectivity/specificity have already been studied and the scope of the method to be validated has already been defined, in terms of matrix types and measurand (for example analyte) concentrations.

2 Normative references

The following document is referred to in the text in such a way that some or all of its content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC Guide 99:2007, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*

3 Terms, definitions and symbols

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC Guide 99 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1.1
measurement
process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity

[SOURCE: ISO/IEC Guide 99:2007, 2.1, modified — Notes to entry have been excluded.]

3.1.2
measurand
quantity intended to be measured

Note 1 to entry: The term “analyte”, employed in chemistry, is a synonym of measurand, and is used more generally.

[SOURCE: ISO/IEC Guide 99:2007, 2.3, modified — Original notes to entry have been excluded and a new note to entry has been added.]

3.1.3
measurement trueness
trueness
closeness of agreement between the average of values obtained by replicate measurements of the same or similar objects under specified conditions and a reference quantity value

[SOURCE: ISO/IEC Guide 99:2007, 2.14, modified — Notes to entry have been excluded.]

3.1.4
measurement precision
precision
closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions

[SOURCE: ISO/IEC Guide 99:2007, 2.15, modified — Notes to entry have been excluded.]

3.1.5
repeatability condition
condition of measurement, out of a set of conditions that includes the same measurement procedure, same operator, same measuring system, same operating conditions and same location, and replicate measurements on the same or similar objects over a short period of time

[SOURCE: ISO/IEC Guide 99:2007, 2.20, modified — Notes to entry have been excluded.]

3.1.6
measurement repeatability
repeatability
measurement precision under a set of *repeatability conditions* (3.1.5) of measurement

[SOURCE: ISO/IEC Guide 99:2007, 2.21]

3.1.7
intermediate precision condition
condition of measurement, out of a set of conditions that includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes

[SOURCE: ISO/IEC Guide 99:2007, 2.22, modified — Notes to entry have been excluded.]

3.1.8
intermediate measurement precision
intermediate precision
measurement precision under a set of *intermediate precision conditions* (3.1.7) of measurement

[SOURCE: ISO/IEC Guide 99:2007, 2.23, modified — Notes to entry have been excluded.]

3.1.9
reproducibility condition of measurement
reproducibility condition
condition of measurement, out of a set of conditions that includes different locations, operators, measuring systems, and replicate measurements on the same or similar objects

[SOURCE: ISO/IEC Guide 99:2007, 2.24, modified — Note to entry has been excluded.]

3.1.10
measurement reproducibility
reproducibility
measurement precision under *reproducibility conditions of measurement* (3.1.9)

[SOURCE: ISO/IEC Guide 99:2007, 2.25, modified — Note to entry has been excluded.]

3.1.11
measurement accuracy
accuracy
closeness of agreement between a measured quantity value and a true quantity value of a measurand

[SOURCE: ISO/IEC Guide 99:2007, 2.13, modified — Notes to entry have been excluded.]

3.1.12
verification
provision of objective evidence that a given item fulfils specified requirements, taking into account any measurement uncertainty

[SOURCE: ISO/IEC Guide 99:2007, 2.44, modified — Notes to entry have been excluded.]

3.1.13
validation
verification, where the specified requirements are adequate for an intended use

Note 1 to entry: The term "characterization" applies to the method, whereas the term "verification" applies to the outcomes. Validation of the method therefore consists of checking if the results are adequate for an intended use.

[SOURCE: ISO/IEC Guide 99:2007, 2.45, modified — Example has been excluded and a Note to entry has been added.]

3.1.14
selectivity
property of a measuring system, used with a specified measurement procedure, whereby it provides measured quantity values for one or more measurands such that the values of each measurand are independent of other measurands or other quantities in the measuring system

Note 1 to entry: The IUPAC considers specificity as the final stage of selectivity.

[SOURCE: ISO/IEC Guide 99:2007, 4.13, modified — Examples and original notes to entry have been excluded. A new note to entry has been added.]

3.1.15
reference value
quantity value whose associated measurement uncertainty is generally considered small enough so that the value may be used as a basis for comparison with quantity values of the same kind

[SOURCE: ISO/IEC Guide 99:2007, 5.18, modified — Notes to entry have been excluded.]

3.1.16
scope
<of the method>all of the types of *matrix* (3.1.22) to which the method applies, taking into account the range of concentrations involved in validation

3.1.17
scope of validation
all of the types of *matrix* (3.1.22) to which the method and range of concentrations involved in validation applies

3.1.18
scope of validity
all of the types of *matrix* (3.1.22) to which the method and range of concentrations involved in validation applies, and for which future outcomes obtained via the method will be considered valid

3.1.19
quantitative method
method of analysis which determines the quantity or weight fraction of an analyte so that it may be expressed as a numeric value in the appropriate units

3.1.20
reference method
method of analysis recognized by experts or used as a reference by agreement between parties, which gives, or is supposed to give the accepted reference value of the measurand

3.1.21
alternative method
method of analysis used by the laboratory instead of one or several *reference methods* (3.1.20)

3.1.22
matrix
set of properties of the sample and its components other than the analyte

Note 1 to entry: The matrix effect reflects the possible influence that these properties or components can have on the instrumental response. For practical reasons, since the matrix effect can vary in the different stages of analysis (e.g. before or after mineralisation), a type of matrix is defined as a group of materials or products recognized by the analyst as having consistent behaviour with regard to the method of analysis used.

3.1.23
series
set of measurements carried out under a set of repeatability conditions

Note 1 to entry: For example, a series includes measurements carried out on the same day and/or by the same operator.

3.1.24
accuracy profile
combination, in a graphic form, of one or several *β-expectation tolerance intervals* (3.1.25) calculated at different concentrations, and of one or several *acceptance intervals* (3.1.26)

3.1.25
β-expectation tolerance interval
tolerance interval
interval which contains, on average, a defined proportion, β %, of future measurements, obtained according to a given procedure and for a given concentration

Note 1 to entry: The limits of the interval are calculated based on trials conducted for the purpose of validation.

Note 2 to entry: A value of 80 % for β % means that, on average, one out of five results will be outside the limits of the interval at the *limit of quantitation* (3.1.29). See 5.10.

3.1.26
acceptance interval
specification of the performance required for the method, expressed as an acceptable deviation around the reference value

Note 1 to entry: The limits of the interval are set by the client or by statutory requirements, sometimes according to the concentration. They are expressed as ±λ as absolute values and in the units of the measurand, or (1 ± λ) × 100 as relative values.

3.1.27
linearity
<of the method> establishment of a linear relationship between the deduced (or quantified) quantities in the samples and their reference values

Note 1 to entry: Linearity of the method is different from linearity of the response function of the measuring apparatus, which only characterizes the instrumental response during calibration and is not essential for accurate quantitation.

3.1.28
validation sample
control sample
material to which the reference value may be assigned, either because it is a reference material (certified or uncertified), or because the molecule to be assayed has been subjected to standard addition

3.1.29
limit of quantitation
the lowest and/or highest concentration of analyte that may be quantified under the experimental conditions of the method. It corresponds to the lowest and/or highest concentration of the *scope of validity* (3.1.18)

Note 1 to entry: Note 1to entry: According to the SFSTP (French society for pharmaceutical sciences and technology), the limit of quantitation is the smallest quantity of analyte in a sample that may be assayed under the experimental conditions described with a defined level of accuracy.

3.1.30
limit of detection
measured quantity value, obtained by a given measurement procedure for which the probability of falsely claiming the absence of a component in a material is b, given a probability, α, of falsely claiming its presence

Note 1 to entry: The notation b used in this definition incurs a risk of type II error.

[SOURCE: ISO/IEC Guide 99:2007, 4.18, modified — Original notes to entry have been omitted and a new note to entry has been added.]

3.2 Symbols

A series of i measurements *(*i varying from 1 to I), includes k concentrations *(*k varying from 1 to K), for which j repetitions have been performed *(*j varying from 1 to J). The subscripts are written in the following order: *i,j,k*. The random variables are written in upper case letters and their values in lower case letters. Description of abbreviations used in formulae is given in Table 1.

Table 1 — Meaning of the different abbreviations used in formulae

Symbol	Description
x_{ijk}	Reference value assigned to a calibration standard for series i ($1 \le i \le I$), repetition j ($1 \le j \le J$) and concentration k ($1 \le k \le K$) or Reference value assigned to a validation sample for series i, repetition j and concentration k.
y_{ijk}	Measurement of the instrumental or experimental response observed for a calibration standard or validation sample for series i, repetition j and concentration k.
z_{ijk}	Deduced value for a validation sample for series i, repetition j and concentration k, obtained either by inverse prediction using a calibration model or by direct measurement.
b_{ijk}	Bias expressing the trueness error for a validation sample between the deduced value and its reference value $b_{ijk} = z_{ijk} - x_{ijk}$

4 General principles

4.1 Reminder

The accuracy profile allows a statistical approach to validation. Formula (1) is used to describe a measurement, z, of a measurand, Z, from a laboratory:

$$z = m + B + e \quad (1)$$

where

m is the overall average for the homogeneous sample sent to the laboratories;

B is the bias component of the laboratory under conditions of repeatability;

e is the random error occurring in each measurement, under conditions of repeatability.

As part of an interlaboratory study, the bias component B comes from the laboratory, but it may also come from any other source of uncertainty in an intralaboratory study, such as the day, operator, instrument, etc.

In addition to the statistical methods for calculating the accuracy criteria, the present document also provides details of the organization of data collection and precautions to be taken.

4.2 Various conditions for the estimation of precision

According to its definition, precision can be estimated under various conditions. In any case, precision is quantified based on a standard deviation, be this for repeatability s_r, intermediate precision s_{IP} or reproducibility s_R. A complexity scale may be established between these different standard deviations, according to the number of sources of uncertainty. Figure 3 illustrates this gradation, from conditions of repeatability where there is no identified variation factor and/or systematic variation component for calculating the deviation between repetitions, to the various possibilities for estimating intermediate precision and, finally, conditions of reproducibility for which the number of sources is not known.

To simplify presentation, the notion of series refers to a set of repetitions performed under conditions of repeatability: a series groups together all of the measurements made under the same conditions, e.g. the same day, the same operator or a short period of time. For certain methods applying to samples that are highly unstable over time, the chosen series effect should be the operator rather than the day; the series will thus include repetitions performed by the same operator:

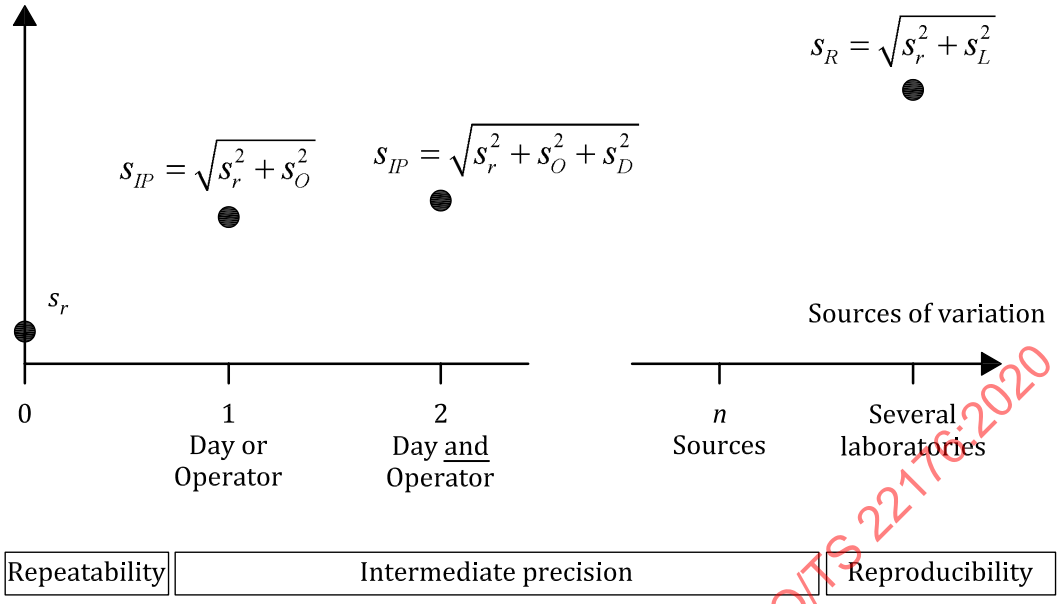


Figure 3 — Various estimations of the precision of a method according to the sources of variation involved

More complicated models may be used, as in the following example, in which different laboratories, days, operators and instruments are combined to give four series in a multi-factorial design with three factors.

Series	Laboratories	Days	Operators	Instruments
1	1	Day 1 (–)	Operator 1 (–)	Instrument 1 (+)
2	1	Day 2 (+)	Operator 1 (–)	Instrument 2 (–)
3	1	Day 1 (–)	Operator 2 (+)	Instrument 2 (–)
4	1	Day 2 (+)	Operator 2 (+)	Instrument 1 (+)

In general, the choice of sources of variation for the measurement series should reflect as best possible the components of variability that are likely to arise upon routine application of the method to be validated.

NOTE For the purposes described in this document, it is essential to collect data in several series and to control the sources of variation. Otherwise, it will not be possible to construct an accuracy profile.

4.3 Accuracy profile

From the intermediate precision or reproducibility standard deviation, calculated according to the calculations described in Annex A, the β-expectation tolerance interval can be obtained, which includes a proportion, β, of future outcomes.

All calculations are performed separately for each concentration k, allowing k precision standard deviations and then k tolerance intervals to be obtained, which are brought together to construct the accuracy profile. Figure 4 shows an example of an accuracy profile constructed using three concentrations, 0,4 mg/L, 2,0 mg/L and 4,0 mg/L, which defines the scope or scope of validation of the method to be validated.

The accuracy profile includes the following graphic elements:

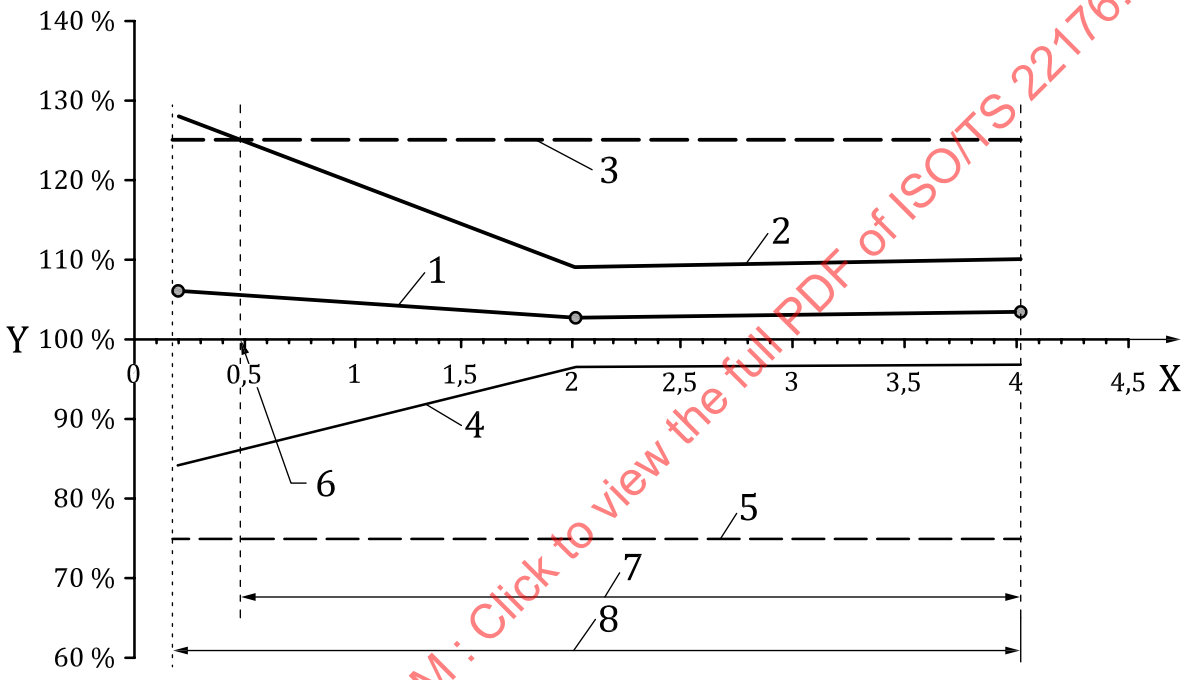
— on the horizontal axis: the theoretical concentrations (the concentration reference values);

— on the vertical axis (simultaneously):

  — the limits of the β–expectation tolerance intervals, calculated from the deduced concentrations and expressed as percentages (recovery rate or relative accuracy);

  — the acceptance intervals, defined according to the method's objective and expressed in the same way as the tolerance intervals.

The interpretation strategy for this graph is described in detail in 5.10. Nevertheless, both the acceptance limits and the proportion β, which is used to calculate the tolerance intervals, are strictly dependent on the method's context of use and shall be adapted to each individual case. In Figure 4, in the area between the broken vertical lines, the method is capable of producing a mean proportion, β, of outcomes that lies within the acceptance limits: the method is therefore valid within this scope. The scope of the method is the scope that was initially chosen for validation.



Key

1 mean recovery rate (%)
2 upper tolerance limit
3 upper acceptance limit
4 lower tolerance limit
5 lower acceptance limit
6 limit of quantification
7 scope of validity
8 scope of validation
X known concentration (%)
Y recovery (mg/L)

Figure 4 — Accuracy profile created using three concentrations

Each grey circle represents the ratio of the mean deduced concentration for the level, expressed as a mean recovery rate (%) for the concentration, and quantifies the trueness. The dotted lines delimit the acceptance interval, and the solid lines delimit the tolerance interval, calculated from the intermediate precision standard deviations for each concentration. The vertical lines delimit the scope of validity, within which the method is capable of producing a high and known percentage of acceptable results.

Generally speaking, in order to comply with the statistical models used, repetitions shall be performed under intermediate precision or reproducibility conditions, using homogeneous validation samples.

5 Procedure

5.1 Definition of the measured quantity

Define the measured quantity according to the procedure for the method, specifying the formulae used to calculate the final result and the method used to reach this result. In the field of cosmetics, we are often in the case of indirect or rational methods, which require prior calibration to calculate the concentration of the unknown samples. These methods involve a two-step approach: firstly, the calibration curve shall be plotted, using the same physicochemical principle used for the samples; measurements are then made using the unknown samples, and their concentrations are calculated using the calibration model. In case of direct methods, which do not require calibration, the situation is even simpler as there is no need of a calibration curve and the measurement on the unknown samples provides the concentration directly.

It is essential that the quantity measured during the trial corresponds to that which will be routinely measured. In particular, if the procedure dictates that the final outcome be expressed following a single measurement, the result of a validation trial shall not be expressed as the mean value of several repetitions.

5.2 Definition of objectives

5.2.1 Choice of the scope of validation

Define the scope of validation of the method in the form of a range of absolute or relative concentrations, e.g. between 1 g and 60 g or between 1 µg/L and 500 µg/L. Practically speaking, it is through the choices made — in terms of the range of concentrations for one or several types of matrix — when developing the experimental design that it is possible to demonstrate the scope within which the method is effectively valid and whether it is able to provide acceptable outcomes (5.4 and 5.5). The scope concerned by this demonstration is called the scope of validity and may be smaller than the previously defined scope of the method.

The choice of the method's scope may correspond to a legal requirement.

If samples are encountered during routine use of the method with concentrations that do not lie within the scope of validity, extrapolation is not permitted. In this case, a complementary study should be carried out to extend the scope, or a dilution may be performed if it can be shown that this has no impact.

Poor definition of the scope of the method can have significant consequences on validity, especially in terms of specificity, interference and cross-reactivity.

5.2.2 Choice of acceptance limits

Acceptance limits are expressed as ± λ, when expressed as absolute values in the same units as the measurand, or (1 ± λ) × 100 as relative values. Wherever possible, define the acceptance limits by referring to a document, a professional practice or a client's requirement.

For information purposes, it could be mentioned that for pharmaceutical products the λ value is usually set at 5 %[11],[12]. However, in cosmetics, due to the complexity and the diversity of the matrices, it could be proposed for ingredient metering (above 0,1 %) a reference value for λ of 10 %. For trace analysis in cosmetics, it could be between 20 % and 40 %[13],[14].

Acceptance limits are generally expressed as percentages. To simplify the presentation of the examples that illustrate this document, a fixed value has been chosen for the entire scope of the method. However, if the scope is wide, values that vary according to the concentration may be used.

5.3 Selection of validation samples

5.3.1 Choice of the type of matrix or types of matrices

The validation samples should be materials that best represent the scope of the method.

Choose a sufficient quantity of stable and homogeneous materials to carry out all of the trials provided in the characterization plan for validation.

5.3.2 Methods for establishing reference values

To estimate the trueness of the method, validation samples should be available whose concentration are known as accurately as possible and with known uncertainty (ISO 11095:1996). This concentration corresponds to the reference value assigned to the validation sample and shall be established independently of the method to be validated. It is expressed as *X*. There are several possible methods for establishing the reference value for a validation sample. These include the following:

1) use certified reference materials (CRMs), external or in-house; the traceability of these materials decreases, respectively;

2) perform standard addition using a standard molecule of known purity. If the reference material already contains endogenous analyte, the experimental design should include a level with no addition so that the measured concentrations may be taken into account in the calculations (see 5.7.3);

3) prepare spiked or synthetic samples. The spiking should be carried out in the earliest stage of the material preparation to account for the analyte-matrix interactions.

To prepare validation samples by standard addition of various concentrations, it is important to reduce dependency between the different preparations as much as possible. Standard additions may be in liquid form (by adding a standard solution) or in solid form (by weighing).

If successive dilutions are prepared using the same stock solution (1/2, 1/4, 1/8, etc.), this creates a correlation between the measurements and exacerbates the consequences of errors in trueness; it is therefore not recommended to use this approach. This note also applies to the preparation of the standard solutions.

The rational choice of reference values is crucial in the application of this document, since it influences the trueness of the method.

5.4 Characterization plan for validation

5.4.1 Organization

The characterization plan is used to estimate the routine performance of the method under routine conditions of implementation.

For this, a trial consists in performing a measurement on a validation sample, with a reference value with known uncertainty. The reference values assigned to concentrations levels *k* can be obtained via the techniques described in 5.1.

The following elements are required for this plan:

— *I* series of measurements ($1 < i < I$);

— for each series, perform *J* repetitions ($1 < j < J$);

— *K* concentrations ($1 < k < K$) to cover the scope of the method.

Fill in a copy of Table 2 with the raw data. In this table, the reference value *X* of the validation samples is expressed in absolute units (mg, g, etc.) or relative units (mg/kg, µg/kg, mg/L, etc.). For indirect

methods, the response Y is expressed in the units corresponding to the instrumental method used (peak area, peak height, absorbance, etc.).

Table 2 — Organization of measurements for the characterization plan for validation

Levels	Series	Concentrations of validation samples (Reference values) X	Measurements (instrumental responses) Y			
			1	2	...	J
1	1	X_{11}	y_{111}	y_{121}	...	y_{1J1}
	...	...	...			
	I		...			
2	1					
	...					
	I					
...	...	...				
K	1					
	...	...				
	I	X_{IK}				y_{IJK}

5.4.2 Choice of the number of series, repetitions and concentrations for the characterization plan for validation

In the context of this document, the minimum requirements for the number of series are:

— a number of series, I, equal to 5, possibly reduced to 4 or 3 if this can be justified. A series may be represented by a particular day, but also by a combination of various sources of uncertainty, such as multiple devices, multiple operators and multiple days;

— a constant number of repetitions per series and per concentration, J, equal to or greater than 2;

— a number of concentrations, K, equal to or greater than 3. It is essential that $K \geq 3$ as this allows the linearity between the concentrations of the reference values and the deduced concentrations to be checked: three concentrations are required for this check. However, when it is necessary to validate the method close to its limit of quantitation LOQ, it is recommended that a K equal to or greater than 4 be chosen.

If $I = 5$, $J = 2$ and $K = 3$, the experimental design will involve 30 tests. Apart from a minimum limit of $I = 3$, a high degree of flexibility is allowed in the choice of the number of tests.

As a reminder, the higher the number of tests, the better the estimates of the validation criteria, i.e. the interval will be narrower and validation easier. However, to improve the results, it is generally preferable to increase the number of series rather than the number of repetitions or concentrations.

5.5 Calibration plan for the indirect methods

5.5.1 Organization

The purpose of the calibration plan is to allow the coefficients of the calibration curve model to be estimated. This plan is only compulsory for indirect methods requiring calibration for quantitation. The experimental design (number of calibration concentrations, repetitions per concentration and the calibration range) should reflect the routine procedure as best possible. The calibration range is not necessarily the same as the scope of the method. To begin with, reagents or standard materials are required, and shall be in one of the forms described in 5.1. For the purpose of the calibration plan, the measurement of a standard of known concentration will be referred to as a "trial".

For this plan, the following is required:

— I series of tests ($1 \leq i \leq I$). Depending on the intermediate precision conditions chosen, the series will consist of tests carried out under repeatability conditions, i.e. on the same day and/or by the same operator and/or using the same instrument. It is essential that this number be the same as for the characterization plan for validation. If the series is represented by a particular day, calibration measurements shall be carried out on the same days as the validation measurements;

— for each concentration, perform J' repetitions ($1 \leq j \leq J'$). J' may be different to J chosen for the characterization plan for validation;

— for each series, K' standard concentrations ($1 \leq k \leq K'$) of known concentration are required, which cover the calibration range.

Draw up a copy of Table 3, which summarizes all of the trials and measurements to be performed. In this table the concentrations of the calibration standards are expressed in absolute units (mg, g, etc.) or relative units (mg/kg, µg/kg, mg/L, etc.).

After performing the trials under conditions of repeatability, the results are noted down in the units given in the instrumental method (peak area, peak height, absorbance, etc.).

Table 3 — Organization of calibration plan trials

Standard levels	Series	Concentrations of calibration samples (reference values) X'	Measurements (instrumental responses) Y'			
			1	2	…	J'
1	1	X'_{11}	y'_{111}	y'_{121}	…	$Y'_{1J'1}$
	…	…	…			
	I		…			
2	1					
	…					
	I					
…	…	…				
K'	1					
	…	…				
	I	$X'_{IK'}$				$y'_{IJ'K'}$

5.5.2 Choice of the number of series, repetitions and concentrations for the calibration plan

The desired value for K' depends on the type of expected instrumental response function, shown in Table 4. The recommended rule is to use at least one more concentration than the number of coefficients in the calibration model. This constraint is related to the use of a least squares regression method to estimate the coefficients of the model (see 5.7.2). However, it is also possible to use the minimum value in the same table. The K' calibration standard concentrations should allow a good statistical estimate of the response function parameters. The least squares method, which is used in the calculations, is very sensitive to the experimental design, and this should therefore be given careful consideration[15],[16].

The number of repetitions J' for each calibration standard in each series is free of choice. However, the recommended number is $J' \geq 2$. During routine use, the standard procedure of the method can be modified according to the conditions that have been proven capable of providing acceptable results.

The values of K' and J' may be different to those of K and J.

If the method requires calibration, it is necessary to ensure that none of the responses, Y, obtained from the validation samples lies outside of the highest or lowest response entered in Table 3. In other words, the calibration model shall not be extrapolated beyond the concentrations for which it has been designed.

5.6 Testing

The measurements described in the characterization plan for validation should be carried out by applying the method as it will be routinely used, ensuring that these measurements are as independent as possible. In particular, the calibration conditions should be applied as best possible, even if they may be modified in light of the results obtained from the accuracy profile. In contrast, the number of "repetitions" used to express a final outcome should be faithfully adhered to. For example, if each final outcome is expressed as the mean of two repetitions, each trial shall be performed according to this method.

The purpose of the trials is to allow the calculation of the method's performance criteria, which will be used to create the accuracy profile and decide upon its validity. The indirect and direct methods require a different number of experimental designs; in the second case a calibration plan is not necessary.

In cases where matrix effects have been revealed during the development of the method, it may be necessary to implement several characterization plans for validation in order to calculate a correction factor that will be routinely applied (except in the case of statutory prohibition). Two characterization plans should thus be implemented using two independent samples, and the results of the first plan used to calculate the correction factor that will be applied to the second plan.

The role of the accuracy profile is to estimate — based on the results obtained — the likelihood that the user will obtain acceptable results when using the method routinely. For this reason, the plans shall be implemented in compliance with the following conditions:

— Intermediate precision or reproducibility conditions. The measurements shall be made under intermediate precision or reproducibility conditions and aim to take into account as many sources of uncertainty as possible. Since these sources of uncertainty depend on the method and the matrices analysed, it is impossible to define a single strategy. It is recommended that a causes and effects diagram be created to choose the most comprehensive conditions.

  In most cases, the measurements may be made over several days, provided that one takes into account the sources of uncertainty (see Annex E) due to the measuring apparatus (instrumental settings, calibration and preparation of reagents), the staff, the preservation of the sample, etc. However, for very unstable samples, measurements may be made over a single day with several operators or instruments.

  NOTE Reproducibility conditions. This document also applies to measurements made under reproducibility conditions and therefore coming from several laboratories.

— Coverage of the scope. By definition, intermediate precision is estimated by repeated measurements using the same object or similar objects, i.e. using a homogeneous sample or several similar samples. A validation study should cover the entire scope of the method. There are several possibilities for dealing with these two constraints:

  — use a single sample whose concentration can be varied, e.g. by means of standard addition (preferred approach for indirect methods);

  — use multiple samples (or even several similar matrices) with various concentrations (typically the case for direct methods);

  — prepare synthetic matrices (raw materials or chemicals);

  — use external or certified reference materials.

In all cases, sufficient quantities of samples for all of the repetitions shall be obtained before beginning.

— Synchronization of the plans. If two plans are necessary (calibration and characterization), it is essential to carry out the trials for the same factors of intermediate precision. For example, if the day is the chosen source of variation, calibration and characterization measurements shall be made on the same day. The calibration data for this day will be used to predict the concentrations from the characterization data of the same day.

5.7 Calculation of predicted inverse concentrations for indirect methods

5.7.1 General

Prior to any numerical processing, it is recommended that an initial graphic illustration and visual examination of the data be carried out in order to detect any blatant errors, such as an incorrectly recorded piece of data. These graphs can be placed in the annex of the assessment file. Abnormal data may be an indication of a poorly-developed method and may nullify validation.

5.7.2 Calculation of the calibration models

For indirect methods, the instrumental response, Y, shall be expressed as a function of the concentrations, x, of the calibration standards, using a mathematical model, f, as shown in Formula (2):

$$Y = f(x) \tag{2}$$

The most typically-used f functions are summarized in Table 4, but this list is not exhaustive. Parameters a_1, a_2, etc. are known as the model parameters.

The parameters of the calibration model from data collected for each i series, are calculated so as to obtain I sets of parameter values. The same type of model shall be used for all of the data, regardless of the series, but the parameters may differ from one series to another. This approach allows any observed inter-series variations to be taken into account.

Table 4 — Main types of possible response functions

Type	Formula	Desirable for K'	Minimum for K'
Line passing through zero	$Y = a_1 x$	2	1
Line	$Y = a_0 + a_1 x$	3	2
Quadratic function	$Y = a_0 + a_1 x + a_2 x^2$	4	3
Logistic function with 4 parameters	$Y = a_0 + \frac{a_3 + a_0}{1 + \left(\frac{a_2}{x}\right)^{a_1}}$	5	5

Calculation of the calibration model coefficient estimates may require various conventional statistical techniques detailed in the reference works:

1) Regression using the least squares method Reference [11]

2) Weighted regression Reference [11]

3) Nonlinear regression Reference [12]

The easiest method is the least squares method, because it is easily available. Weighted regression applies in cases where the response variances are not homogeneous between concentrations. The model parameter estimates for each series are summarized in Table 5.

If the type of calibration model was clearly established during the development of the method, the minimum number for K' may be used. However, it is best to use the desirable number, or more, in order to check whether a more complex model may be more suitable.

It can be noted that from the data of the calibration plan, it is possible to estimate several calibration models and thus create multiple accuracy profiles. The model which provides the most favourable profile can then be selected. It is then necessary to modify the procedure according to the model chosen, without having to change the acceptance limits λ.

NOTE For direct methods, this calculation is not possible.

Table 5 — Model parameter estimates according to the series

Series	1	...	*I*
a_0			
a_1			

5.7.3 Calculation of back-calculated concentrations by inverse prediction

For indirect methods, the calibration models (see Table 4) are used to calculate the deduced concentrations from the data of the validation plan by using the inverse function of the calibration model, according to the Formula (3):

$$\hat{x} = z = f^{1}(Y) \tag{3}$$

The inverse function is called the inverse prediction formula, and the values thus obtained, z, are called deduced concentrations. Table 6 shows these formulae for each of the response functions in Table 4.

Use the data of the validation plan and the parameters of the calibration models estimated from the calibration plan to calculate the deduced concentrations of the validation samples. Perform the calculations series by series, with the model for the corresponding series. This method of calculation underlines the fact that it is essential for the trials of the calibration and validation plans to be performed for the same series *i*, that is on the same day and/or by the same operator and/or by the same laboratory (see 4.2).

Table 6 — Inverse prediction equations according to the type of response function

Response function	Reverse function for back calculated concentration calculation
Line forced through zero	$z = \frac{Y}{a_1}$
Line	$z = \frac{Y - a_0}{a_1}$
Quadratic function	$z = \frac{-a_1 + \sqrt{a^2{}_1 - 4a_2(a_0 - Y)}}{2a_2}$
Logistic function with 4 parameters	$z = \frac{a_2}{\left(\frac{a_3 - a_0}{Y - a_0} - 1\right)^{\frac{1}{a_1}}}$

5.8 Calculation of the validation criteria by concentration level

5.8.1 General

Present the raw data and the final outcomes (deduced concentrations) as described in Annex B and illustrated in Annex C on a true example.

5.8.2 Trueness criteria by series

Summarize all of the data in a copy of Table 7. In this table, it is possible to include one or several trueness criteria (in the form of bias). These criteria are:

Absolute bias as shown in Formula (4):

$$b_{ijk} = Z_{ijk} - X_{ijk} \tag{4}$$

Relative bias as shown in Formula (5):

$$b_{ijk}\,\% = \frac{Z_{ijk} - X_{ijk}}{X_{ijk}} \times 100 \tag{5}$$

Recovery as shown in Formula (6):

$$Rc = \frac{Z_{ijk}}{X_k} \times 100 \tag{6}$$

Table 7 — Concentrations deduced by inverse prediction from the data of the validation plan

Level	Series	Repetition	Theoretical concentration *X*	Deduced concentration *Z*	Absolute bias	Relative bias
1	1	1	x_{111}	z_{111}	b_{1jk}	b_{1jk} %
1	1	2				
1		…				
…		*J*				
1	*I*	1				
1	*I*	2				
1	*I*	…				
1	*I*	*J*				
K	*I*	*J*	x_{IJK}	z_{IJK}	b_{IJK}	b_{IJK} %

If different calibration models are tested, a table is required for each model.

Use the "*z*" data from Table 7 to calculate the repeatability, inter-series and intermediate precision (or reproducibility) standard deviations. This calculation is performed independently for each concentration *k* (where $1 \le k \le K$), as described in Annex A.

5.8.3 Trueness and precision criteria by concentration

Calculate the precision criteria, as described in Annex A. Summarize these together with the trueness criteria in a copy of Table 8.

Table 8 — Precision and trueness criteria by concentration

Criteria	Symbol	Levels		
		1	…	*K*
Mean reference value	$\overline{x}_k$			
Mean back-calculated concentration	$\overline{z}_k$			
Repeatability standard deviation	s_{kr}			
Inter-series standard deviation	s_{kB}			
Intermediate precision standard deviation	$s_{kIP} = \sqrt{s_{kr}^2 + s_{kB}^2}$			

Table 8 *(continued)*

Criteria	Symbol	Levels		
		1	...	*K*
Intermediate coefficient of variation	$\frac{s_{kIP}}{\overline{x}_k}\times 100$			
Mean absolute bias	$\overline{z}_k - \overline{x}_k$			
Mean relative bias	$\left(\frac{\overline{z}_k - \overline{x}_k}{\overline{x}_k}\right)\times 100$			
Mean recovery rate	$\frac{\overline{z}_k}{\overline{x}_k}\times 100$			

5.8.4 Calculation of the tolerance intervals

For the purpose of this document, the method of calculation proposed by Mee[10] was chosen. It was also adopted by a committee of the French society for pharmaceutical sciences and technology (SFSTP) [1][17]. The calculation is performed using the data in Table 8, and independently for each concentration, *k*. For simplification, the subscript *k* has been omitted from Formulae (7) to (11).

Express the tolerance interval as a symmetric interval around the mean calculated concentration, *z*, as shown in Formula (7):

$$z \pm k_{tol} \times s_{TI} \tag{7}$$

Calculate the standard deviation of the tolerance interval s_{TI} using Formula (8).

$$s_{TI} = s_{IP}\left(\sqrt{1+\frac{1}{I\times J\times B^2}}\right) \tag{8}$$

$$B = \sqrt{\frac{R+1}{J\times R+1}} \tag{9}$$

where $R = \frac{s_B^2}{s_r^2}$

Quantity k_{tol} is the coverage factor of the tolerance interval and is obtained using Formula (10):

$$k_{tol} = t_{\upsilon,\frac{1+\beta}{2}} \tag{10}$$

where

$t_{\upsilon,\frac{1+\beta}{2}}$ is the quantile of Student's *t* distribution with υ degrees of freedom;

β is the expected probability of the content of the tolerance interval.

The number of degrees of freedom, ν, is calculated using the Satterthwaite approximation[4], shown in Formula (11).

$$\upsilon = \frac{(R+1)^2}{\frac{\left(R+\frac{1}{J}\right)^2}{I-1} + \frac{1-\frac{1}{J}}{IJ}} \tag{11}$$

For each concentration, k, the standard deviation of the tolerance interval, s_{TI}, the coverage factor, k_{tol}, and the lower and upper limits of the tolerance interval are calculated. All of the calculations are summarized in a copy of Table 9.

Table 9 — Tolerance interval limits by concentration

Criteria	Symbol	Levels		
		1	...	K
Mean reference value	$\bar{x}_k$			
Low tolerance limit	$\bar{z}_k - k_{tol} \times s_{TI}$			
High tolerance limit	$\bar{z}_k + k_{tol} \times s_{TI}$			
Relative low tolerance limit	$\left(\frac{\bar{z}_k - k_{tol} \times s_{TI}}{\bar{x}}\right) \times 100$			
Relative high tolerance limit	$\left(\frac{\bar{z}_k + k_{tol} \times s_{TI}}{\bar{x}}\right) \times 100$			
Relative low acceptance value	$(1 - \lambda) \times 100$			
Relative high acceptance value	$(1 + \lambda) \times 100$			

It can be noted that ν (see Formula 11) is rarely an integer, and it is necessary to use Student tables that include fractional values of degrees of freedom. However, it is possible to approach the quantile value via linear interpolation between the two whole degree of freedom values surrounding ν.

NOTE The ratio $R = \frac{s_B^2}{s_r^2}$ between the inter-series variance and the repeatability variance intervenes in several formulae. It reflects the relative importance of the series effect. For example, in the case where the day effect is the series effect, if the method is very stable from one day to another and is capable of providing very similar results for the same sample, R is close to 1 and will only play a secondary role. However, if this ratio increases, the number of degrees of freedom decreases, and the smaller the number of degrees of freedom, the higher the quantile of Student's t distribution and the wider the tolerance interval.

The parameter β represents the probability of obtaining outcomes outside of the acceptance limits at the LQ. Beyond the LQ, the probability of obtaining unacceptable outcomes is much lower than β.

5.9 Construction of the accuracy profile

The accuracy profile may be constructed in different ways, according to the type of data being processed. The most conventional method when dealing with relative concentrations is that presented in Figure 4 in which performances are expressed relatively, as a recovery rate.

To construct the accuracy profile, the following data are required:

a) on the horizontal axis:

— the mean reference values.

b) on the vertical axis:

— the lower relative tolerance limits;

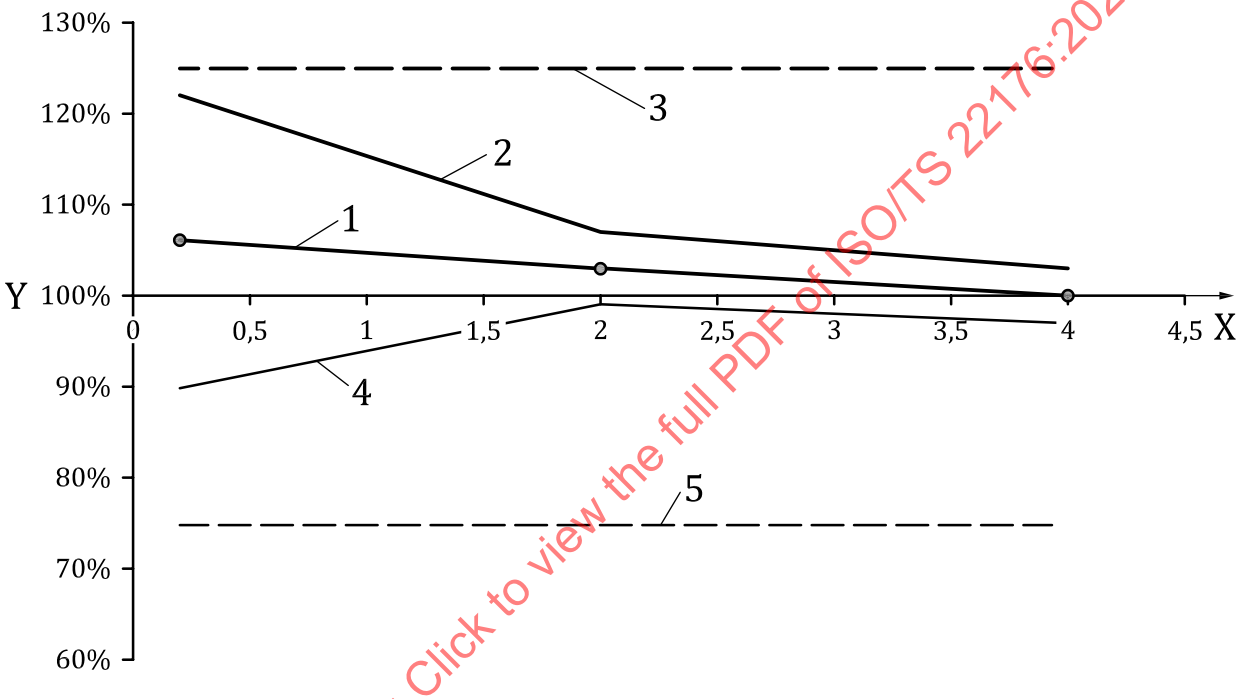
— the upper relative tolerance limits;

— the mean recovery rates;

— the lower relative acceptance limits;

— the upper relative acceptance limits.

These data are plotted on a graph, with the reference values on the horizontal axis, as shown in Figure 5.



Key

1 mean recovery rate
2 upper tolerance limit
3 upper acceptance limit
4 lower tolerance limit
5 lower acceptance limit
Y recovery (%)
X concentration

Figure 5 — Accuracy profile expressed by the recovery rate

5.10 Interpretation of the accuracy profile for validation

5.10.1 General

To use the accuracy profile to validate a method, the following two decision criteria shall be set:

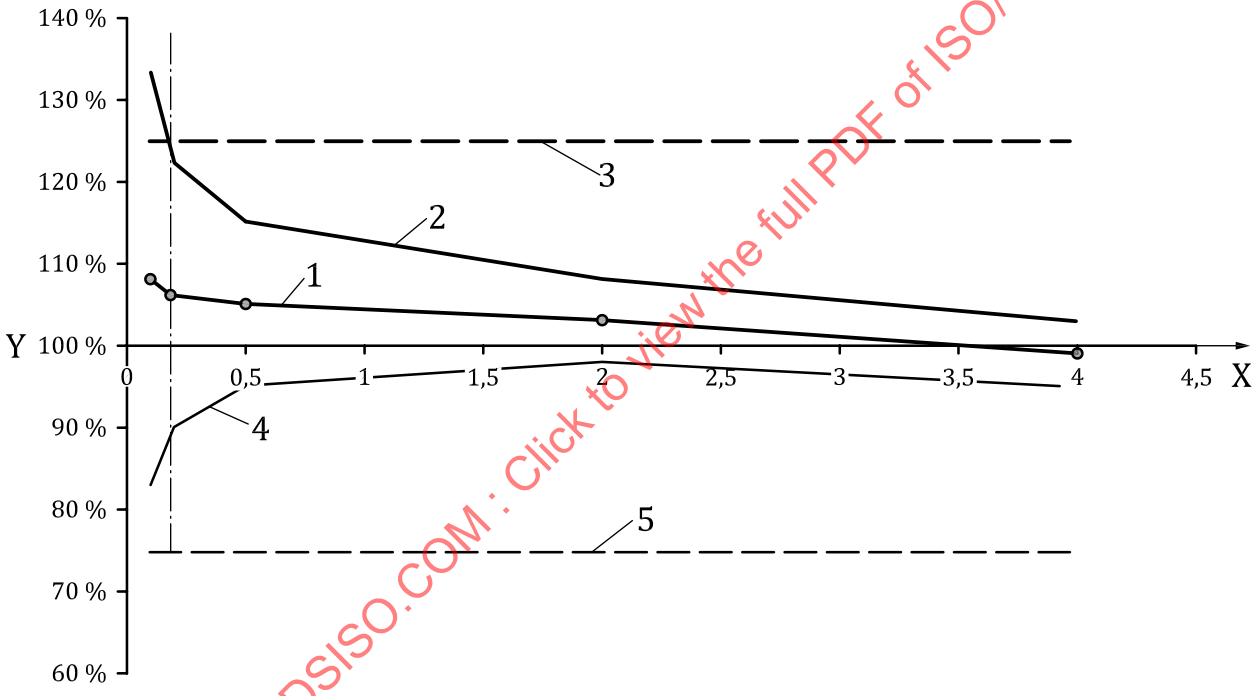
— **Acceptance limits.** These reflect the practical objectives of the users. They delimit an interval either side of the reference value. These limits are usually regulatory requirements. However, if there is no set reference value, the expectations of the end users should be considered, such as a given limit of quantitation (see 5.2.2.);

— **β proportion.** This represents the proportion of future outcomes that will lie, on average, within the tolerance intervals. Based on the β values found in the literature for different domains, the recommended value for β in cosmetics is 80 %. It means that, in average, 4 out of 5 experimental results will fall inside the tolerance interval.

5.10.2 Decision rules

Plot the acceptance limits on the accuracy profile graph to allow direct visual interpretation of the results. Figure 6 illustrates a typical situation in which most of the tolerance intervals fall within the acceptance interval. If the tolerance interval does not lie within the acceptance interval, the method is considered unable to provide a sufficient number of acceptable outcomes based on the choices made at the start of the study. For example, for β = 80 % and an acceptance limit of ± 25 %, the method may be considered valid between about 0,2 mg/kg and 4,0 mg/kg.

This graph also gives other indications. In particular, in Figure 6, it can be seen that trueness varies with concentration. The recovery rate which reflects trueness is around 108 % at low concentrations and is close to 100 % at high concentrations. Various elements explaining the behaviour of the method's bias can be found using the statistical results in Table 8.

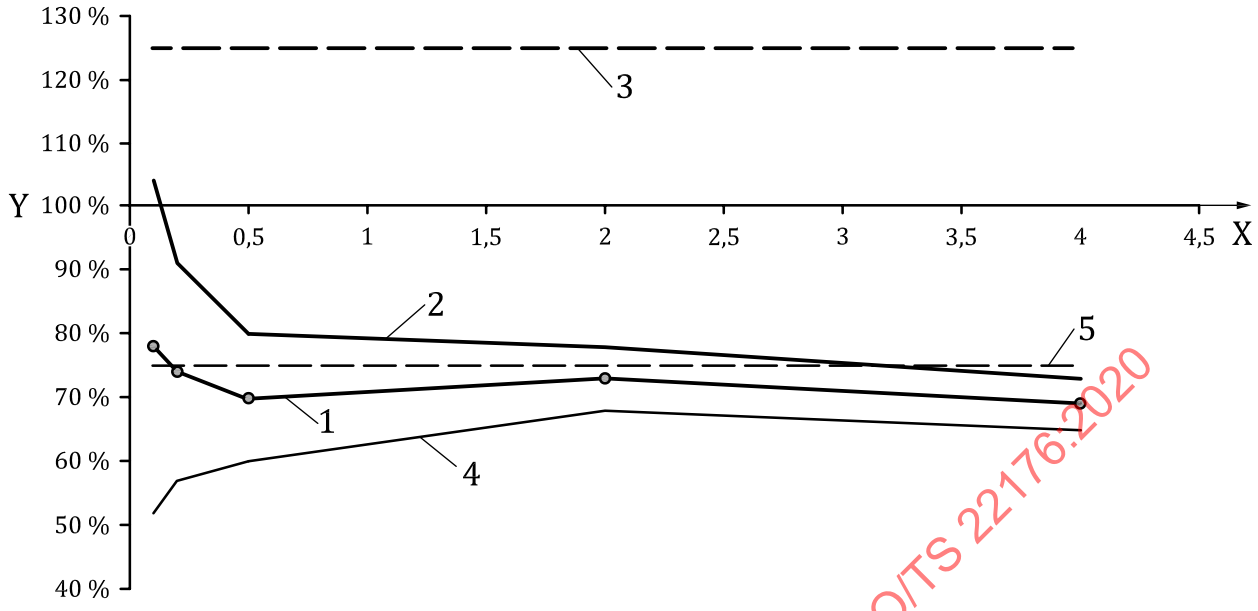


Key

1 mean recovery rate
2 upper tolerance limit
3 upper acceptance limit
4 lower tolerance limit
5 lower acceptance limit
Y recovery (%)
X concentration (mg/kg)

Figure 6 — Example of a validated method for a limited area of scope of validation

In the case illustrated in Figure 7, the method has high systematic bias - the recovery rate is around 70 % - which may be due, for example, to an uncontrolled matrix effect. It should therefore be concluded that the method is not valid within the scope of the study. Professional practices and regulations permitting, a correction factor may be applied to correct this bias.



Key

1 mean recovery rate
2 upper tolerance limit
3 upper acceptance limit
4 lower tolerance limit
5 lower acceptance limit
Y recovery (%)
X concentration (mg/kg)

Figure 7 — Example of an accuracy profile for a method having a high systematic bias, characterized by a recovery rate around 70 %

5.10.3 Definition of the scope of validity

The scope of validity is determined by the area of the scope of validation within which the method provides an acceptable percentage of results, at least equal to β. It is delimited by a lower boundary that corresponds to the lower limit of quantitation, and an upper boundary that corresponds to the upper limit of quantitation.

5.10.4 Choice of a calibration procedure for the routine

Provided, for indirect methods, that calibration data are available, it is possible to create several accuracy profiles with the same set of data, but using different calibration models, e.g. a model running through zero based on a single standard solution can be compared with a complete linear model calculated from three standard solutions. It is thus possible to select either the most favourable profile or the simplest calibration procedure that allows the set objective to be achieved.

5.10.5 Influence and significance of the β proportion

As described in 5.10.1, the parameter β represents the proportion of future outcomes that are expected to lie within the tolerance intervals. The risk of error does not produce inaccurate results. It is rather the risk that the analyst takes of having to repeat a full analysis because the method does not produce an acceptable result for a sample of known concentration. A value of 80 % means that the risk of this occurring is, on average, one in five. Of course, a higher value reduces the risk, but makes for a more stringent validation process. Annex D shows the influence of β.

5.10.6 Identification of outliers

The method of calculation proposed in this document requires that the experimental design for validation be balanced (see Annex A). That is why it does not provide for any methods of detecting and/or rejecting atypical data values (outliers).

Nevertheless, if the person in charge of validation considers the rejection of outliers necessary following a visual examination of a graph, he/she is free to implement the procedure that he/she feels best suited. It is also always possible to do again measurements for a given concentration or data series.

NOTE A high number of outliers indicates poor implementation of the method of analysis and can compromise its validation.

6 Management of the outcomes during routine use

During routine application of the method thus validated, it is necessary to regularly check that the results obtained remain acceptable. This can be done, for example, by means of a control chart, in particular if a correction factor is used.

Ineffective corrective measures may also result in a need to revalidate the method. Revalidation may be complete or partial; partial revalidation requires an experimental design for a single concentration to check whether the tolerance interval obtained remains between the acceptance limits.

Annex A
(normative)

Calculation of repeatability, intermediate precision and reproducibility standard deviations

A.1 General

Calculations shall be carried out independently for each concentration *k,* using the concentrations deduced by inverse prediction for indirect methods or the deduced concentrations for direct methods. In both cases, the deduced concentrations are written Z_{ijk}. By convention, the subscript i ($1 \le i \le I$) represents the number of the series and j ($1 \le j \le J$) the number of the repetition in the series. To simplify notation, the subscript k ($1 \le k \le K$) for the concentration is omitted in the following formulae.

For each concentration k, calculate the following:

1) the repeatability standard deviation, written s_r;

2) the inter-series standard deviation, written s_B;

3) the intermediate precision standard deviation, $s_{IP} = \sqrt{S_r^2 + S_B^2}$.

A.2 Case of a balanced plan

It takes into account experimental designs for which the number of repetitions per series, I, is the same for all of the series. The total number of repetitions shall be equal to IJ. However, this total may be different for each concentration.

Calculate the overall mean for a concentration using Formula (A.1):

$$\overline{\overline{Z}} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} Z_{ij}}{I \times J} \tag{A.1}$$

Calculate the repeatability, inter-series and intermediate precision standard deviations using a random effects model analysis of variance (ANOVA), according to the principles described in ISO 5725-2:1994, namely by decomposition of the total sum of squared deviations into two sums of squared deviations:

$$\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(z_{ij} - \overline{\overline{Z}}\right)^2}{SSD_t} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(z_{ij} - \overline{\overline{Z}}\right)^2}{SSD_r} + \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(z_{ij} - \overline{\overline{Z}}\right)^2}{SSD_B} \tag{A.2}$$

Formula (A.2) is conventionally written in an abbreviated form, involving three sums of squared deviations *(SSDs).*

$SSD_t = SSD_B + SSD_r$ General analysis of variance equation,

for which each of the sums is defined as follows in order to facilitate interpretation:

— SSD_t Total sum of deviations from the overall mean for the concentration;

— SSD_B Sum of inter-series deviations;

— SSD_r Sum of intra-series deviations.

It is not necessary to develop the three sums of squares in order to perform the calculations; SSD_B is calculated by subtraction. This method can be problematic if the result is negative. If this is the case, impose a value of 0 for SSD_B.

$$SSD_B = SSD_t - SSD_r \text{ if } SSD_B > 0$$

$$SSD_B = 0 \text{ if } SSD_B \leq 0$$

Calculate the repeatability variance for the concentration from the repetitions Z_{ij}, as shown in Formula (A.3):

$$S_r^2 = \frac{SSD_B}{I(J-1)} \tag{A.3}$$

Calculate the inter-series variance — written S_B^2 — as follows:

$$S_B^2 = \frac{\frac{SSD_B}{I-1} - S_r^2}{J} \tag{A.4}$$

Finally, calculate the intermediate precision standard deviation for the concentration:

$$S_{IP} = S_B^2 + S_r^2 \tag{A.5}$$

See Annex C for an example of calculation using Excel.

A.3 Case of an unbalanced plan

Since this document does not cover tests for the rejection of atypical data values, but rather leaves users the freedom to disregard values that they consider atypical, this may result in an unbalanced experimental design. In this case, this alternative algorithm is proposed in order to avoid the loss of any data.

If the number of repetitions n_i is not the same for all of the series for a particular concentration, use Formula (A.6) instead of Formula (A.4) to calculate the inter-series variance:

$$S_B^2 = \frac{(I-1)\left(\frac{SSD_B}{I-1} - S_r^2\right)}{N^*} \tag{A.6}$$

where

$$N^* = N - \frac{\sum_{i=1}^{I} n_i^2}{N}$$

$$N = \sum_{i=1}^{I} n_i$$

Use the conventional formulae for the other precision criteria.

NOTE An alternative to this calculation method is to use a variance estimation algorithm that uses the restricted maximum likelihood method (REML); this is available in most specialized statistical software programs.

Annex B
(normative)

Contents of the validation file

B.1 Documentary part

Describe the procedure in enough detail for any competent person to be able to use it (including the calculations). Draft it according to the following outline, which is not exhaustive:

— title;

— warnings and safety precautions;

— introduction;

— purpose and scope;

— normative references (if applicable) and bibliographical references;

— definitions;

— principle;

— reagents and products;

— equipment;

— sampling and sample preparation;

— procedure;

— expression of results;

— specific cases;

— notes;

— test report;

— bibliography;

— appendices.

B.2 Experimental part

a) Choice of the scope of acceptance and acceptance limits

b) Validation materials and setting of target values

c) Characterization plan for validation

d) Calibration plan (for indirect methods)

e) Estimation of calibration model coefficients (for indirect methods)

f) Calculation of predicted inverse concentrations and calculation of trueness

g) Calculation of precision data and tolerance intervals

h) Construction of the accuracy profile

i) Interpretation

Table B.1 — Summary of performance criteria

Criteria	Concentration A	Concentration B	Concentration C	Related section
Upper acceptance limit				5.2.2
Lower acceptance limit				5.2.2
Reference value (X),				5.3.2
Deduced mean value (Z),				5.8.2 and 5.8.3
Trueness (X-Z),				5.8.2 and 5.8.3
Repeatability standard deviation				Annex A
Intermediate precision standard deviation				Annex A
Coefficient of variation of intermediate precision (%)				
Repeatability limit				

Table B.2 — Other useful criteria

Other criteria	Related section
Limit of quantitation (LQ)	
Limit of detection (LD)	
Specificity (Yes/No)	
Mean recovery rate (%)	5.8.3

j) Conclusions

Annex C
(informative)

Setting-up an assay for determining the accuracy profile in the case of NDELA in cosmetic samples

C.1 General

The measurement of N-nitrosodiethanolamine (NDELA) in cosmetics has been already described in two International Standards (see ISO 15819 and ISO 10130). In order to complete the validation of the methods, it is then an opportunity to describe here what could be done with NDELA using the accuracy profile approach. This study has been carried out according to ISO 15819 in a single laboratory in intermediate precision conditions.

— To follow the validation procedure described in Clause 5, below are the different items set up.

Determination of an accuracy profile for method ISO 15819 to quantify NDELA (see 5.1) in one cosmetic matrix (see 5.3.1) between 25 and 400 µg/kg of NDELA (see 5.2.1).

As this is an indirect analytical method (see 5.1) and according to its description in the ISO 15819:2014 document, calibration curve was performed by using a line response function (see 5.7.2).

Acceptance limits were set at λ = ± 20 % (see 5.2.2) and tolerance intervals will be calculated with a proportion β = 80 % (see 5.10).

Validation samples were prepared by spiking a same cosmetic matrix with NDELA (see 5.3.2) at different concentration levels between 25 and 400 µg/kg.

I = 5 series were considered and differed by the day (see 5.4.1).

For each *i* series, measurements were collected the same day (see 5.6) as follows:

— for calibration standards with K' = 5 and J' = 1 (see 5.5.1), see Table C.1 for details of K';

— for validation samples with K = 4 and J = 4 (see 5.4.1), see Table C.2 for details of K.

NOTE 1 Deuterated NDELA d8 is used in every calibration solutions and validation samples as internal standard.

NOTE 2 In order to minimize the potential impact of other matrix components that might prevent NDELA quantitation to prevent the measuring instrument getting dirty and to ensure that none of the measurements *Y* obtained from the validation samples lie outside of the highest or lowest response of calibration solutions, two kinds of assay sample solutions were done depending on each validation sample. For VS 1 and 2, assay sample solutions were diluted at 10 % in order to detect a concentration of NDELA equivalent to 2,5 µg/kg and 5 µg/kg respectively. For VS 3 and 4, assay sample solutions were diluted at 2 % in order to detect a concentration of NDELA equivalent to 3 µg/kg and 8 µg/kg respectively. In every case, attention was paid to also introduce the internal standard NDELA d8 at 20 µg/kg.

Table C.1 — Levels of calibration standards (equivalent to Table 3)

Levels of CS (k')	Series (i)	Concentrations of CS (Reference values) $X'_{ij'k'}$			Measurements (instrumental responses) $Y'_{ij'k'}$
			NDELA	NDELA d8	
CS 1	1	X'_{111}	1,01 µg/kg	20 µg/kg	y'_{111}
	2	X'_{211}	1,01 µg/kg	20 µg/kg	y'_{211}
	3	X'_{311}	1,01 µg/kg	20 µg/kg	y'_{311}
	4	X'_{411}	1,01 µg/kg	20 µg/kg	y'_{411}
	5	X'_{511}	1,01 µg/kg	20 µg/kg	y'_{511}
CS 2	1	X'_{112}	2,02 µg/kg	20 µg/kg	y'_{112}
	2	X'_{212}	2,02 µg/kg	20 µg/kg	y'_{212}
	3	X'_{312}	2,02 µg/kg	20 µg/kg	y'_{312}
	4	X'_{412}	2,02 µg/kg	20 µg/kg	y'_{412}
	5	X'_{512}	2,02 µg/kg	20 µg/kg	y'_{512}
CS 3	1	X'_{113}	5,06 µg/kg	20 µg/kg	y'_{113}
	2	X'_{213}	5,06 µg/kg	20 µg/kg	y'_{213}
	3	X'_{313}	5,06 µg/kg	20 µg/kg	y'_{313}
	4	X'_{413}	5,06 µg/kg	20 µg/kg	y'_{413}
	5	X'_{513}	5,06 µg/kg	20 µg/kg	y'_{513}
CS 4	1	X'_{114}	10,12 µg/kg	20 µg/kg	y'_{114}
	2	X'_{214}	10,12 µg/kg	20 µg/kg	y'_{214}
	3	X'_{314}	10,12 µg/kg	20 µg/kg	y'_{314}
	4	X'_{414}	10,12 µg/kg	20 µg/kg	y'_{414}
	5	X'_{514}	10,12 µg/kg	20 µg/kg	y'_{514}
CS 5	1	X'_{115}	20,24 µg/kg	20 µg/kg	y'_{115}
	2	X'_{215}	20,24 µg/kg	20 µg/kg	y'_{215}
	3	X'_{315}	20,24 µg/kg	20 µg/kg	y'_{315}
	4	X'_{415}	20,24 µg/kg	20 µg/kg	y'_{415}
	5	X'_{515}	20,24 µg/kg	20 µg/kg	y'_{515}

Table C.2 — levels of validation samples (equivalent to Table 2)

Levels of VS (k)	Series (i)	NDELA concentrations of VS (Reference values) X_{ijk}	Measurements (instrumental responses) Y_{ijk}
VS 1	1	$X_{111} = X_{121} = X_{131} = X_{141} = 23{,}4$ µg/kg	$y_{111} = y_{121} = y_{131} = y_{141}$
	2	$X_{211} = X_{221} = X_{231} = X_{241} = 23{,}4$ µg/kg	$y_{211} = y_{221} = y_{231} = y_{241}$
	3	$X_{311} = X_{321} = X_{331} = X_{341} = 23{,}4$ µg/kg	$y_{311} = y_{321} = y_{331} = y_{341}$
	4	$X_{411} = X_{421} = X_{431} = X_{441} = 23{,}4$ µg/kg	$y_{411} = y_{421} = y_{431} = y_{441}$
	5	$X_{511} = X_{521} = X_{531} = X_{541} = 23{,}4$ µg/kg	$y_{511} = y_{521} = y_{531} = y_{541}$
VS 2	1	$X_{112} = X_{122} = X_{132} = X_{142} = 46{,}7$ µg/kg	$y_{112} = y_{122} = y_{132} = y_{142}$
	2	$X_{212} = X_{222} = X_{232} = X_{242} = 46{,}7$ µg/kg	$y_{212} = y_{222} = y_{232} = y_{242}$
	3	$X_{312} = X_{322} = X_{332} = X_{342} = 46{,}7$ µg/kg	$y_{312} = y_{322} = y_{332} = y_{342}$
	4	$X_{412} = X_{422} = X_{432} = X_{442} = 46{,}7$ µg/kg	$y_{412} = y_{422} = y_{432} = y_{442}$
	5	$X_{512} = X_{522} = X_{532} = X_{542} = 46{,}7$ µg/kg	$y_{512} = y_{522} = y_{532} = y_{542}$
VS 3	1	$X_{113} = X_{123} = X_{133} = X_{143} = 146{,}1$ µg/kg	$y_{113} = y_{123} = y_{133} = y_{143}$
	2	$X_{213} = X_{223} = X_{233} = X_{243} = 146{,}1$ µg/kg	$y_{213} = y_{223} = y_{233} = y_{243}$

Table C.2 *(continued)*

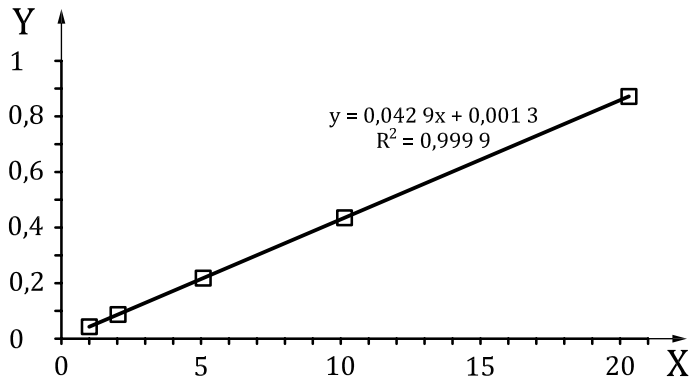
Levels of VS (k)	Series (i)	NDELA concentrations of VS (Reference values) X_{ijk}	Measurements (instrumental responses) Y_{ijk}
	3	$X_{313} = X_{323} = X_{333} = X_{343}$ = 146,1 µg/kg	$y_{313} = y_{323} = y_{333} = y_{343}$
	4	$X_{413} = X_{423} = X_{433} = X_{443}$ = 146,1 µg/kg	$y_{413} = y_{423} = y_{433} = y_{443}$
	5	$X_{513} = X_{523} = X_{533} = X_{543}$ = 146,1 µg/kg	$y_{513} = y_{523} = y_{533} = y_{543}$
VS 4	1	$X_{114} = X_{124} = X_{134} = X_{144}$ = 389,7 µg/kg	$y_{114} = y_{124} = y_{134} = y_{144}$
	2	$X_{214} = X_{224} = X_{234} = X_{244}$ = 389,7 µg/kg	$y_{214} = y_{224} = y_{234} = y_{244}$
	3	$X_{314} = X_{324} = X_{334} = X_{344}$ = 389,7 µg/kg	$y_{314} = y_{324} = y_{334} = y_{344}$
	4	$X_{414} = X_{424} = X_{434} = X_{444}$ = 389,7 µg/kg	$y_{414} = y_{424} = y_{434} = y_{444}$
	5	$X_{514} = X_{524} = X_{534} = X_{544}$ = 389,7 µg/kg	$y_{514} = y_{524} = y_{534} = y_{544}$

C.2 Calculation of predicted inverse concentrations (5.7)

1) For each i series, regression parameters were determined according to a line response function (see 5.7.2) and summarized in Table C.3. Data were also checked with a graphical illustration (see example for Series 1 in Figure C.1).

Table C.3 — Measured values for calibration standards and corresponding regression parameters (equivalent to Table 5)

Calibration standard	Concentrations $X'_{ij'k'}$	Measurements $Y'_{ij'k'}$				
		Series/Day 1	Series Day 2	Series Day 3	Serie Day 4	Series/Day 5
CS 1	1,01 µg/kg	0,048	0,040	0,052	0,051	0,053
CS 2	2,02 µg/kg	0,085	0,089	0,090	0,101	0,103
CS 3	5,06 µg/kg	0,215	0,234	0,218	0,246	0,222
CS 4	10,12 µg/kg	0,438	0,485	0,463	0,462	0,460
CS 5	20,24 µg/kg	0,868	0,967	1,014	0,910	0,859
Slope a_1		0,042 9	0,048 2	0,050 3	0,044 4	0,042 0
y-intercept a_0		0,001 3	−0,007 9	−0,019 1	0,012 3	0,016 0
R^2		0,999 9	0,999 9	0,997 2	0,999 8	0,998 9



Key

Y Y measurement

X NDELA concentration (µg/kg)

Figure C.1 — Graphical illustration of line response function for calibration standards of Series 1

2) Then for each *i* series, back-calculated concentrations of validation samples were deduced by using inverse prediction equation indicated in Table 6 and corresponding regression parameters of the series (see 5.7.3).

NOTE For VS 1 and 2, the calculation took into account the real assay concentration at about 10 %, and for VS 3 and 4, the calculation took into account the real assay concentration at about 2 %:

$$\text{For VS 1 and 2: } z_{ijk} = \frac{Y_{ijk} - a_{i0}}{a_{i1}} \times \frac{100}{C_{assay}} \text{ with } C_{assay} \approx 10\,\%$$

$$\text{For VS 3 and 4: } z_{ijk} = \frac{Y_{ijk} - a_{i0}}{a_{i1}} \times \frac{100}{C_{assay}} \text{ with } C_{assay} \approx 2\,\%$$

where C_{assay} is assay concentration of sample, expressed as a percentage (a mass fraction).

See example for Series 1 in Table C.4 and all deduced concentrations for all series in Table C.5.

Table C.4 — Calculations done for Series 1 to deduce concentrations for validation samples

Validation samples	Repetition # J	Assay sample conc.	Measurements Y_{ijk}	VS deduced conc. Z_{ijk}
VS 1	1	10,38 %	0,114	25,3 µg/kg
	2	10,12 %	0,096	21,8 µg/kg
	3	10,54 %	0,105	22,9 µg/kg
	4	10,28 %	0,100	22,4 µg/kg
VS 2	1	10,48 %	0,225	49,8 µg/kg
	2	10,20 %	0,179	40,7 µg/kg
	3	10,20 %	0,188	42,8 µg/kg
	4	10,06 %	0,209	48,2 µg/kg
VS 3	1	2,22 %	0,138	143,4 µg/kg
	2	2,29 %	0,134	135,6 µg/kg
	3	2,32 %	0,131	130,6 µg/kg
	4	2,26 %	0,144	147,1 µg/kg

Table C.4 *(continued)*

Validation samples	Repetition # J	Assay sample conc.	Measurements Y_{ijk}	VS deduced conc. Z_{ijk}
VS 4	1	2,43 %	0,428	409,3 µg/kg
	2	2,31 %	0,405	407,9 µg/kg
	3	2,11 %	0,349	385,1 µg/kg
	4	2,12 %	0,322	353,4 µg/kg

Table C.5 — Deduced concentrations for validation samples and their corresponding absolute bias and relative bias (equivalent to Table 7)

Column	A	B	C	D	E	F	G
Line							
1	Level (k)	Series (i)	Repetition (j)	Theoritical concentration X (µg/kg)	Deduced concentration Z_{ijk}	Absolute bias b_{ijk}	Relative bias b_{ijk} %
2	1	1	1	23,4	25,3	1,9	8,1 %
3	1	1	2	23,4	21,8	−1,6	−6,6 %
4	1	1	3	23,4	22,9	−0,5	−2,2 %
5	1	1	4	23,4	22,4	−1,0	−4,4 %
6	1	2	1	23,4	21,2	−2,2	−9,6 %
7	1	2	2	23,4	20,0	−3,4	−14,5 %
8	1	2	3	23,4	26,1	2,7	11,4 %
9	1	2	4	23,4	23,9	0,5	2,1 %
10	1	3	1	23,4	23,5	0,1	0,5 %
11	1	3	2	23,4	24,9	1,5	6,4 %
12	1	3	3	23,4	25,1	1,7	7,5 %
13	1	3	4	23,4	22,0	−1,4	−5,8 %
14	1	4	1	23,4	25,3	1,9	8,0 %
15	1	4	2	23,4	25,8	2,4	10,3 %
16	1	4	3	23,4	19,5	−3,9	−16,8 %
17	1	4	4	23,4	27,9	4,5	19,2 %
18	1	5	1	23,4	27,5	4,1	17,6 %
19	1	5	2	23,4	26,3	2,9	12,3 %
20	1	5	3	23,4	25,4	2,0	8,5 %
21	1	5	4	23,4	23,4	0,0	0,1 %
22	2	1	1	46,7	49,8	3,1	6,7 %
23	2	1	2	46,7	40,7	−6,0	−12,9 %
24	2	1	3	46,7	42,8	−3,9	−8,3 %
25	2	1	4	46,7	48,2	1,5	3,1 %
26	2	2	1	46,7	45,7	−1,0	−2,2 %
27	2	2	2	46,7	42,0	−4,7	−10,1 %
28	2	2	3	46,7	44,3	−2,4	−5,1 %
29	2	2	4	46,7	47,5	0,8	1,7 %
30	2	3	1	46,7	42,9	−3,8	−8,1 %
31	2	3	2	46,7	43,5	−3,2	−6,8 %
32	2	3	3	46,7	48,9	2,2	4,7 %

Table C.5 *(continued)*

Column	A	B	C	D	E	F	G
Line							
1	Level (k)	Series (i)	Repetition (j)	Theoritical concentration X (µg/kg)	Deduced concentration Z_{ijk}	Absolute bias b_{ijk}	Relative bias b_{ijk} %
33	2	3	4	46,7	48,2	1,5	3,3 %
34	2	4	1	46,7	51,9	5,2	11,2 %
35	2	4	2	46,7	42,8	−3,9	−8,3 %
36	2	4	3	46,7	54,3	7,6	16,4 %
37	2	4	4	46,7	50,6	3,9	8,4 %
38	2	5	1	46,7	44,8	−1,9	−4,1 %
39	2	5	2	46,7	41,5	−5,2	−11,2 %
40	2	5	3	46,7	46,1	−0,6	−1,2 %
41	2	5	4	46,7	49,8	3,1	6,6 %
42	3	1	1	146,1	143,4	−2,7	−1,8 %
43	3	1	2	146,1	135,6	−10,5	−7,2 %
44	3	1	3	146,1	130,6	−15,5	−10,6 %
45	3	1	4	146,1	147,1	1,0	0,7 %
46	3	2	1	146,1	136,2	−9,9	−6,8 %
47	3	2	2	146,1	133,6	−12,5	−8,6 %
48	3	2	3	146,1	126,0	−20,1	−13,8 %
49	3	2	4	146,1	115,0	−31,1	−21,3 %
50	3	3	1	146,1	122,4	−23,7	−16,2 %
51	3	3	2	146,1	131,6	−14,5	−9,9 %
52	3	3	3	146,1	146,3	0,2	0,1 %
53	3	3	4	146,1	143,5	−2,6	−1,8 %
54	3	4	1	146,1	118,0	−28,1	−19,2 %
55	3	4	2	146,1	129,0	−17,1	−11,7 %
56	3	4	3	146,1	138,7	−7,4	−5,1 %
57	3	4	4	146,1	124,9	−21,2	−14,5 %
58	3	5	1	146,1	142,3	−3,8	−2,6 %
59	3	5	2	146,1	149,3	3,2	2,2 %
60	3	5	3	146,1	137,1	−9,0	−6,2 %
61	3	5	4	146,1	159,2	13,1	9,0 %
62	4	1	1	389,7	409,3	19,6	5,0 %
63	4	1	2	389,7	407,9	18,2	4,7 %
64	4	1	3	389,7	385,1	−4,6	−1,2 %
65	4	1	4	389,7	353,4	−36,3	−9,3 %
66	4	2	1	389,7	361,2	−28,5	−7,3 %
67	4	2	2	389,7	348,8	−40,9	−10,5 %
68	4	2	3	389,7	340,6	−49,1	−12,6 %
69	4	2	4	389,7	316,1	−73,6	−18,9 %
70	4	3	1	389,7	341,5	−48,2	−12,4 %
71	4	3	2	389,7	431,0	41,3	10,6 %
72	4	3	3	389,7	351,3	−38,4	−9,8 %