

Guide to Method Validation for Quantitative Analysis in Chemical Testing Laboratories (ISO 17025)

PS15

TABLE OF CONTENTS

1.	Foreword	2
2.	Scope	3
3.	References	4
4.	Definitions	6
5.	Scale of method validation	12
6.	Review of performance characteristics for a method of analysis	17
7.	Follow-up of method validation	26
8	Laboratory competence	26
9	Laboratory documentation	27
10	INAB Contact	27

1. FOREWORD

With the introduction of EN ISO/IEC 17025, the requirements governing the documentation of methods, including method selection and validation of methods, have been amplified. The level of documentation required by INAB in the field of chemical testing still depends on whether standard methods, modified methods or a laboratory's own methods are involved.

As there is scope for interpretation in any standard, it has long been desirable from many quarters to have guidance drawn up to ensure more uniform interpretation of the requirements for validation.

In drawing up this document, it has proved necessary not only to interpret the requirements of the standard with respect to laboratories in the sector but also to provide guidance on the assessment of laboratories. This document is generally to be regarded as a guideline for chemical analysis laboratories, but also contains a few sections that are purely aimed at INAB assessors in this area. Where sections are purely aimed at INAB assessors, this is apparent from the text. In addition, the document contains instructions and examples of how method validation can be carried out.

2. SCOPE

INAB guidelines concern interpretation of relevant sections of the accreditation criteria, drawn up with a view to ensuring consistent case administration and assessment of applicant and accredited laboratories and to place the treatment of Irish laboratories on a par with laboratories in other countries.

This guideline lays down INAB's interpretation of point 5.4 and in particular point 5.4.5 of EN ISO/IEC 17025:1999 (as revised by ISO/IEC 17025:2005) for chemical analysis laboratories, concerning the laboratories' validation of methods of analysis adopted and to the assessment of the results of such validation work. The guideline has been drawn up on the basis of current international standards and guidelines in the sector.

Measurement uncertainty (uncertainty budgets), which should be part of method validation, is not dealt with at length in this document. Reference is in this respect made to the GUM document [15] and the Eurachem/CITAC document [16].

In this issue, the guideline covers exclusively quantitative, analytical methods; however, it does not cover sampling in connection with the performance of these methods.

Where specific validation requirements are laid down by Regulatory Authorities (in legislation, etc.) compliance with the specified validation protocols will be accepted by INAB as demonstration of fitness for purpose.

A few sections purely describe guidance for INAB's assessors and the document is therefore also called Internal Guidance. In addition, the document contains guidance and examples of how method validation can be performed (section 6). The guideline cannot, however, as such be regarded as a manual for method validation in connection with the laboratories' compliance with the requirements of ISO 17025.

The guideline applies to all applicant and accredited chemical analysis testing laboratories.

3. REFERENCES

- [1] McNaught & Wilkinson: *Compendium of Chemical Terminology*, 2nd ed., IUPAC, 1997 (ISBN 0-86542-684-8).
Also known as IUPAC's "Gold Book"
Available on: www.iupac.org/publications/compendium/index.html
- [2] *International vocabulary of basic and general terms in metrology* (VIM) (BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, 2nd ed., 1993).
- [3] ISO/IEC 17025: *General Requirements for the competence of testing and calibration laboratories*, 2nd edition, May 2005.
- [4] CITAC/Eurachem Guide: *Guide to Quality in Analytical Chemistry, An Aid to Accreditation*. Prepared jointly by CITAC (The Cooperation on International Traceability in Analytical Chemistry) and Eurachem (A Focus for Analytical Chemistry in Europe), 2002.
a) The guide is based on the previous documents 'CITAC Guide 1' and 'Eurachem Guidance Document No. 1/WELAC Guidance Document No. WGD 2', with account being taken of the change from EN 45001 to ISO/IEC 17025 as accreditation standard.
b) The guide is based in particular on the CITAC document, with relatively few amendments and additions being made.
c) The previous document has been abolished as an official EA document (previously EA-4/05); this document has (for the time being) not been granted equivalent status.
- [5] IUPAC: *Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC Technical report)*. Resulting from the Symposium on Harmonization of Quality Assurance Systems for Analytical Laboratories, Budapest, Hungary, 4-5 November 1999, held under the sponsorship of IUPAC, ISO & AOAC INTERNATIONAL. Prepared for publication by Michael Thompson, Stephen L.R. Ellison and Roger Wood. Published in Pure Appl. Chem. Vol. 74, No. 5, 2002 (pp. 835 – 855).
- [6] IUPAC: *Harmonized guidelines for the use of recovery information in analytical measurement (Technical report)*. Resulting from the symposium on harmonization of internal quality assurance systems for analytical laboratories, Orlando DC, USA, 4-5 September 1996, held under the sponsorship of IUPAC, ISO & AOAC International.
- [7] Eurachem Guide: *The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics*, 1st edition, Eurachem, Dec. 1998.
- [8] EN ISO 9000: *Quality control systems – Basic principles and glossary*. .
- [9] Youden, W.J. & Steiner, E.H.: *Statistical Manual of the AOAC*, Association of Official Analytical Chemists, 1975.
- [10] ISO 5725-1: *Accuracy (trueness and precision) of measurement methods and results. Part 1: General principles and definitions*, 1995.
- [11] ISO 3534-1:1993: *Statistics – Vocabulary and symbols. Part 1: Probability and general statistical terms*, ISO, 1993.
- [12] ISO Guide 30:1992, *Terms and definitions used in connection with reference materials*, ISO 1992.
- [13]

[14]

- [15] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML: *Guide to the expression of uncertainty in measurement*, ISO, Geneva, 1993.

This document is often referred to as "GUM". It is interpreted in the chemical, analytical sector in the form of the document *Quantifying Uncertainty in Analytical Measurement* [16].

- [16] Eurachem, CITAC: *Quantifying Uncertainty in Analytical Measurement*, 2nd edition, 2000

This document is designated by Eurachem and CITAC as QUAM:2000 P1.

EA refers to the document as an official interpretation of the accreditation requirements.

- [17] Poulsen, Holst & Molin Christensen: *Method evaluation and quality control of chemical measurement processes*, The National Institute of Occupational Health, Copenhagen 1993. (ISBN 87-7534-440-8).

- [18] Funk, Dammann & Donnevert: *Quality Assurance in Analytical Chemistry*, VCH Verlagsgesellschaft mbH, Weinheim, Germany, 1995 (ISBN 3-527-28668).

- [19] IUPAC: *Nomenclature in evaluation of analytical methods including detection and quantitation capabilities. Recommendations*, 1995. Published in Pure Appl. Chem. Vol. 67 (1995) pp. 1699 – 1723.

- [20] Directive 96/23/EC Measure to Monitor certain Substances and Residues thereof in Live Animals and Animal Products

- [21] INAB Policy on Flexible Scopes, PS11

4. DEFINITIONS

The terms used in this guideline are defined here in order to ensure a clear and unambiguous understanding of the terms. Where possible, internationally accepted definitions have been taken as the starting point; however, there are cases where these are not clear and so a definition of our own has been chosen for this guideline.

Many of the terms may also have other – more or less sector-specific – designations, which are acceptable if the laboratory has made the meaning clear within the context of validation.

It should be pointed out that notes connected with the definition in the quoted reference (NOTE(S)) have been included, where possible. In some cases, certain comments important for understanding the definition in question in relation to this guideline have also been added (COMMENT(S)).

4.1 Method of Analysis

Method describing in detail a procedure of chemical analysis with a view to determining/identifying one or more analytical parameters.

COMMENT:

- a) “Method of analysis” is the term generally accepted and adopted in laboratories, but there is no official definition of the term.
- b) “Measurement method” is the logical sequence of operations, generically described, used for the performance of measurements.
- c) “Measurement procedure” is the set of operations, described in detail, used in the performance of certain measurements in accordance with a given measurement method.

4.2 Analytical Parameter / Analyte

The element, chemical compound or property to be detected and quantified by the performance of a method of analysis.

COMMENT:

In this guideline, the term “analytical parameter” is used not only for purely chemical substances and compounds but also for magnitudes such as dry matter, pH, conductivity, various collective parameters, etc.

4.3 Bias [10]

The difference between the calculated mean of the measurement results and an accepted reference value.

NOTE:

Bias is the total systematic error, as opposed to random error. There may be one or more systematic error components contributing to bias. A substantial departure from the accepted reference value is manifested in a high level of bias.

4.4 Blank Value

A blank value is obtained as a result of analysis of a specimen which does not, as far as possible, contain the analytical parameter(s) in question - and with the implementation of the method of analysis.

COMMENTS:

- a) The quoted type of blank value, also called “test blank”, is the magnitude referred to in the rest of this guideline.
- b) Unlike a “test blank”, a “reagent blank” is also adopted in certain contexts [see 4.30].

4.5 Limit of Detection

There are at least two different interpretations in relation to the definition of the term "limit of detection":

1. The limit of detection defined on the basis of the desire to avoid a specimen not containing analytical parameters erroneously being ascribed a concentration (false positive), and at the same time avoid a specimen containing the analytical parameter being erroneously reported as "undetected" (false negative). The definition of the limit of detection then becomes the lower content of an analytical parameter that can be detected with a specifically defined level of confidence (e.g. 95, 99 or 99.7%) using a given method of analysis.
2. The limit of detection defined on the basis of a desire for a certain maximum uncertainty in the result. The definition of the limit of detection then becomes the minimum content that can be measured with a reasonable statistical certainty [9].

COMMENTS:

- a) Expressed in statistical terms, definition 1 above can also be formulated as "*The result of a measurement performed with a given measurement procedure, for which the probability of an analytical false negative is β , given that the probability of an analytical false positive is α .*" IUPAC recommends that α and β should in principle be assigned the value 0.05.
- b) If a measurement gives a value below the level of detection, the result can be quoted as "less than the limit of detection" – NOT as "zero".

4.6 Degrees of Freedom

The number of independent determinations (estimates) of a given statistical magnitude (e.g. mean or standard deviation) that can be performed on the basis of a given data set.

COMMENT:

The greater the number of degrees of freedom, the better the statistical basis for the determination of the magnitude in question.

4.7 Sensitivity [2,9]

The change in response on a measuring instrument divided by the corresponding change in stimulus.

NOTE:

Stimulus may for example be the amount of analyte present. Sensitivity may depend on the value for the stimulus in question. Although this definition clearly concerns measuring instruments, it can also be used in connection with methods of analysis as a whole, with account also being taken of other factors such as specimen preparation, etc.

COMMENT:

Sensitivity can often be expressed by the gradient of the linear calibration curve.

4.8 Performance Parameter/Characteristic

Characteristics for a method of analysis determining its possibilities and limitations (measuring range, specificity, precision, etc.)

4.9 Trueness [10]

Level of conformity between the mean for a population of measurement results and the accepted reference value.

NOTE:

Trueness is normally expressed as the bias of the method.

4.10 Limit of Quantitation

The lowest concentration that can be determined with acceptable laboratory reproducibility and trueness.

COMMENT:

The limit of quantitation forms the lower level for the measuring range.

4.11 Laboratory Reproducibility

Precision under reproducible conditions, within the laboratory.

4.12 Linearity

The ability to induce a signal (response) that is directly proportional to the concentration of the given analytical parameter.

COMMENTS:

- a) Determination of linearity is used among other things in connection with the formulation of a calibration curve.
- b) For some methods, calibrations may give non-linear fits e.g. quadratic fit

4.13 Linear Range [1]

Concentration range within which the intensity of the signal obtained is directly proportional to the concentration of the analytical parameter inducing the signal.

COMMENT:

The original text [1] uses the slightly specific expression “species” for what has here been translated as “analytical parameters/analyte”.

4.14 Matrix [1]

All components in the specimen, with the exception of the analytical parameter(s) in question.

4.15 Measuring Range

The concentration range within which the analytical parameter in question can be determined with specifically determined trueness and precision.

NOTES:

- a) The measuring range may certainly contain a broader concentration range than a linear range, with the link between response and concentration not in all cases being regarded as linear.
- b) The measuring range must be defined in relation to the analyte, matrix and relevant range of use.
- c) Trueness and precision must be acceptable throughout the validated measuring range.

4.16 Measurement Uncertainty [2]

A parameter linked to the measurement result and characterising the spread of the values that can reasonably be added to the measurement magnitude.

NOTES:

- a) This parameter may, for example, be a standard deviation (or a given multiple thereof) or half the breadth of an interval that has a quoted confidence level.
- b) Measurement uncertainty generally comprises many components. Some of these can be assessed on the basis of the statistical distribution of the results of measurement series and can be characterised by experimental standard deviations. Other components, which are also characterised by standard deviations, are assessed on the basis of assumed probability distributions based on experience or on other information.
- c) It should be understood that the measurement result is the best estimate of the value of the measurement magnitude and that all uncertainty components, including those originating from systematic effects (e.g. components associated with corrections and reference standards), contribute to the spread.

[This definition is the same as in the “Guide to the expression of uncertainty in measurement” [15] in which the logical reasons are detailed (particularly in section 2.2.4 and Annex D)].

4.17 Accuracy [10]

Degree of conformity between a measurement result and the accepted reference value.

NOTE:

Used in connection with a range of measurement results, the expression “accuracy” covers a combination of random components and a common systematic error or bias component

COMMENT:

In GUM [15], the definition is expressed in relation to the “true value of the measurement magnitude” instead of “... the accepted reference value”.

4.18 Precision [10]

Degree of conformity between independent measurement results obtained under prescribed conditions.

NOTES:

- a) Precision depends only on the distribution of the random error, and is unrelated to a true value, a conventional true value or an accepted reference value.
- b) Precision is often expressed as a “lack of precision” and is calculated as the spread of the measurement results. Less precision corresponds to greater spread.
- c) The phrase “independent measurement results” means that the results are obtained in a manner such that they are unaffected by any previous results for the same or similar measurement objects. Quantitative measures for precision depend critically on the prescribed measurement conditions. Repeatability and reproducibility conditions are extreme points for prescribed measurement conditions.

4.19 Repeatability [10]

Precision under repeatable conditions.

4.20 Repeatable Conditions [10]

Conditions in which independent measurement results are obtained by the same method on identical specimens in the same laboratory, by the same operator, with the same equipment and within short intervals of time.

4.21 Repeatability Standard Deviation [10]

The spread in the distribution of measurement results obtained under repeatable conditions.

NOTES:

- a) Repeatability standard deviation is an expression of the variation in the distribution of measurement results under repeatable conditions.
- b) Accordingly, “repeatability variance” and “repeatability variation coefficient” could be defined and used as an expression of the variation of measurement results under repeatable conditions.

4.22 Reproducibility [10]

Precision under reproducible conditions.

4.23 Reproducible Conditions [10]

Conditions under which independent measurement results are obtained with the same method on identical specimens in different laboratories by different operators with different equipment.

4.24 Reproducibility Standard Deviation [10]

The spread in the distribution of measurement results obtained under reproducible conditions.

NOTES:

- a) Reproducibility standard deviation is an expression of the variation in the distribution of measurement results under reproducible conditions.
- b) Accordingly, “reproducibility variance” and “reproducibility variation coefficient” could be defined as an expression of the variation in the distribution of measurement results under reproducible conditions.

4.25 Robustness/Ruggedness [8]

The degree of independence of the method of analysis from minor deviations in the experimental conditions of the method of analysis.

4.26 Selectivity [4]

The selectivity of a method of analysis refers to the degree to which the method of analysis is usable for determining the presence of specific analytical parameters in a complex mixture (matrix) without interference from other analytical parameters in the mix.

COMMENTS:

- a) The ability of a method of analysis to measure only what it is intended to measure (the “specificity” of the method).
- b) A lack of selectivity in a method of analysis is manifested, among other things, in the **interference** which can hamper determination of the analytical parameter sought.
- c) Selectivity should be viewed in relation to the analytical parameter in question – whether this is an individual substance or a group of substances.

4.27 Validation [8]

Confirmation – via the provision of objective evidence* - that the requirements for specifically intended use or application have been met.

NOTES:

- a) The expression “validated” is used to denote the corresponding status.
- b) In the case of validation, the use/application conditions may be actual or simulated.
- c) *Objective evidence is defined in [8] as data substantiating the presence or veracity of something (Objective evidence can be furnished by observation, measurement, testing or in other ways).

COMMENT:

The difference between the terms “validation” and “verification” in terms of their practical use is discussed at greater length in section 5 of this guideline.

4.28 Recovery

Detection of a known amount of an analytical parameter added to a specimen and included throughout the method of analysis. After deducting any detected content of the analytical parameter in question in the original specimen without addition, the **recovery percentage** can be calculated as a percentage of the amount added.

COMMENTS:

- a) The aim must be for the addition to be at a level in proportion to the content of the specimen, so different dilutions need not be performed, but also so large that the difference between the content of the specimen and the specimen plus added analytical parameter can be determined with an uncertainty no greater than for determination of the content of the specimen.
- b) In validation, recovery tests are used to monitor the overall effect of different procedure steps, for example evaporation losses, extraction, clean-up, but should not be used purely to assess any one of the effects in isolation.
- c) Recovery tests are used to assess bias.

4.29 Reagent Blank

Reagent blank determination means the complete analytical procedure applied without the test portion or using an equivalent amount of suitable solvent in place of the test portion [20].

5 SCALE OF METHOD VALIDATION

The laboratory must, in connection with any testing within the accreditation area, have undertaken a conscious assessment of the fitness of the method of analysis used for the purpose of the given test. Fitness must be adapted to the applicant's needs and be documented by appropriate validation, with a suitable validation programme (a protocol) being set up and implemented for each matrix type for which the method of analysis is intended to be used.

There are a number of general, overriding requirements concerning how extensive such a validation programme should be (included in ISO/IEC 17025). It is also apparent from this, however, that the scale of a validation programme as a whole and the scale of the assessment of the individual performance characteristics must be laid down in relation to the specific situation.

This section of the guideline examines a number of important aspects concerning definition of the relevant scale of method validation. The individual performance characteristics and their importance in connection with method validation are examined in section 6 of this guideline.

5.1 Validation/Verification

The terms 'validation' and 'verification' are often linked when referring to the necessary scale of method validation, with verification often being regarded as a process that is less extensive (and thus less demanding) than validation.

INAB has developed a separate policy on verification of test methods which is documented in PS24.

It is important that the laboratory has undertaken an assessment of needs for the specific task and, based on this, laid down a validation programme (a protocol), considering

- 1) the nature of the task – including the applicant's requirements in this respect
- 2) the method of analysis in question and the documentation of this in relation to a task of this nature

If new requirements concerning the method of analysis have not been laid down in this situation and all relevant performance characteristics have been fully documented in connection with previous validation, the laboratory may choose to verify one or more (where appropriate, all) performance characteristics. The laboratory must in this respect have documented the fact that, during performance of the method of analysis in the given situation, the specifications laid down in connection with previous validation work are observed. This previous validation work may be performed by the laboratory itself or originate from publication and documentation of an accepted method.

Depending on the situation, validation work may thus consist of anything from full validation of all performance characteristics (i.e. no verifications) to only verifications of all relevant performance characteristics (see the summary in section 5.3).

In cases where the laboratory can justify verification only of a previously validated performance characteristic, the requirements in PS24 shall be followed,

Any use of verification as part of validation work must be well-founded – particularly if the laboratory has accreditation for a "flexible scope" (see INAB Policy Statement PS11), with importance being attached to the laboratory's competence and routines in the area in question.

5.2 General Requirements Concerning the Scope of Method Validation

Validation must be as extensive as necessary to meet the requirements in connection with the given use or the given application.

Validation should (as a minimum) encompass:

- 1) a specification of the applicant's need in connection with the testing in question;

- 2) a definition of the performance characteristics of the method of analysis in relation to the respective requirement;
- 3) an investigation of whether the needs can be met by using the method of analysis with the performance characteristics laid down;
- 4) a declaration concerning the validity of the method of analysis.

In the resulting declaration, there must be justification as to why some of the performance characteristics of the method of analysis may not have been found relevant in connection with the validation in question. This might for instance be omission of validation of the limit of detection if the method is exclusively to be used at high concentrations.

If there are sector-specific requirements in relation to the form and scope of a method validation, the laboratory shall follow these requirements.

5.3 Scope of Method Validation Dependent on Type of Method of Analysis

The scope of method validation required shall be assessed for each specific analyte and shall establish the methodology to be used.

In the accreditation standard, ISO/IEC 17025, it is stated in point 5.4.2 that:

"The laboratory shall use test and/or calibration methods, including methods for sampling, which meet the needs of the customer and which are appropriate for the tests and/or calibrations it undertakes preferably those published as international, regional or national standards" and further: "When the customer does not specify the method to be used, the laboratory shall select appropriate methods that have been published either in international, regional or national standards, or by reputable technical organizations, or in relevant scientific texts or journals, or as specified by the manufacturer. Laboratory-developed methods or methods adopted by the laboratory may also be used if they are appropriate for the intended use and if they are validated. The customer shall be informed as to the method chosen.

The laboratory shall confirm that it can properly operate standardized methods before introducing the tests or calibrations. If the standard method changes, the confirmation shall be repeated".

Standardised methods of analysis are published by standardisation bodies (e.g. CEN, ISO) and recognised technical organisations (e.g. NMKL, AOAC, etc.), etc., but the documentation for the standardised method of analysis varies. In some cases there is no validation report as a basis for the method of analysis, or the performance characteristics are not – or only partially – validated. If this is the case, verification of the laboratory's ability to use the method of analysis is not directly possible, and validation is necessary. In this guideline, the expression "accepted method of analysis" is also used at several points for methods of analysis of the above nature, with there therefore being no certainty as to how well such a method of analysis is documented.

5.4 Situation-Dependent Requirements Concerning the Scope of Method Validation

Some examples that can be used as the starting point for assessment in various situations are set out below. In general, it should be mentioned that limits of quantitation only need to be included if this is required for the specific testing field.

Examples of the scale of method validation/verification in a number of different situations:

1. The method of analysis is developed internally. This includes use of a standardised method of analysis on a matrix which is not described in the method	Completely internal validation: the laboratory must give consideration to all performance characteristics (see section 6), and perform appropriate validation of them. In the event of omission of certain performance characteristics in the validation work, the laboratory shall account for the omission.
2a The method of analysis is published in the scientific literature, but lacks important performance characteristics	Full internal validation as indicated under point 1.
2b The method of analysis is published in the scientific literature and specifies important performance characteristics	The laboratory shall verify its ability to use the method of analysis with respect to the specified performance characteristics. The laboratory must assess the need for validation of any performance characteristics.
3a The method of analysis is standardised or otherwise well-established, but data from external validation is unavailable	The laboratory shall assess the need for validation of those performance characteristics, which are critical for the method of analysis. The laboratory must as a minimum document its ability to use the method of analysis with quality that meets the customer's needs.
3b The method of analysis is published by one of the standardisation organisations, e.g., S.I., CEN, AOAC, NMKL, ISO, etc. The method of analysis is externally validated	The laboratory shall verify its ability to meet the specified performance characteristics for the method of analysis. Selectivity and robustness can reasonably be expected to be investigated in connection with the standardisation process.

4. The method of analysis is well-established within the testing field (cf. point 2b or 3b) but is used with modification	The laboratory shall undertake and document an assessment of what possible effects modification of the method of analysis may have on the individual performance characteristics. In cases where a relevant effect is involved, the performance characteristics in question shall be validated. The laboratory shall verify the modified method against the specifications for performance characteristics in the original method.
5. Change of method of analysis	The scale is in principle the same as indicated under point 2 and 3. If the laboratory's existing method of analysis for validation/verification proves to be free of known systematic error, the procedure for investigating trueness may be a comparison of the results obtained with the new and the old method of analysis. The possibility of investigating trueness is thus improved as any natural specimen can be included in the work.
6. Extension of the measuring range of a method of analysis already commissioned and validated	Trueness and precision shall be investigated for the new measuring range, e.g. in the case of investigation of homogeneous inspection materials (point 5.5) or participation in inter-laboratory investigations.
7. Change from equipment to new equipment but with the same measurement principle	Measuring range, linearity, limit of detection, trueness and precision shall be compared with data from the old equipment with the aid of homogeneous inspection materials (verification) (see section 5.5).
8. Extension of range of matrix types	Trueness and precision (performance characteristics) shall be investigated for the new matrix type, as is deemed fit for purpose.

5.5 Use of Homogeneous Inspection Materials

In this guideline, reference is repeatedly made to the use of “homogeneous inspection materials” as the basis for laying down the trueness, precision, etc. of the method of analysis. Depending on whether traceability is to be documented in the given phase of the validation work, the laboratory may in principle choose between several types (levels) of inspection materials. When determining trueness, the highest level of traceability of the inspection material shall always be sought.

The homogeneous inspection materials have traditionally been described as follows (ascending level of traceability):

1. *Inspection material at user level* – i.e. according to experience, stable and homogeneous test material (not necessarily monitored systematically for stability and homogeneity) with one or more properties that have a well-defined composition that corresponds to the specimen's. Inspection materials are suitable for the ongoing monitoring of equipment or the performance characteristics of the method of analysis to guarantee and document compliance with prescribed quality requirements. They are used only to determine repeatability and reproducibility and not for calibration.
2. *Commercially available inspection material*, which is monitored for homogeneity and stability, where appropriate with a specified assigned value which is not determined by a procedure warranting the description of ‘reference material’. Such materials may be used in internal quality control instead of the materials described above. In addition, various laboratories may compare analysis results for these materials, i.e. comparative tests between laboratories without traceability. Such materials are often described as reference materials, which, according to international standards, is not correct.
3. *Reference materials* are stable and homogeneous materials with one or more properties that are well-defined to be used for calibration of equipment, assessment of a method of analysis or to guarantee the traceability of inspection materials.
4. *Certified reference materials* are reference materials with one or more properties whose values are determined with the aid of technically valid measurement procedures, and which are accompanied by – or traceable to – a certificate or other documentation.

These certified reference materials may be used for calibrating equipment, for validating and monitoring methods of analysis, for monitoring the competence of laboratory personnel or when two or more methods of analysis are to be compared.

The definitions of 3 and 4 are in accordance with the official definitions in ISO Guide 30 [12]. The laboratory should have considered the use of homogeneous inspection materials at a relevant level in connection with their validation work, with the need to use inspection material with a relevant matrix and traceability requirements being catered for as far as possible.

5.6 Assessment of Scale

INAB assessors shall:

1. assess whether the laboratory has adequately covered requirements for the method of analysis (e.g. official requirements or other generally applicable requirements in connection with routine analyses; if new or revised methods of analysis are involved in connection with the extension of accreditation under “flexible scopes”). The laboratory must be able to document close dialogue with the customer with a view to determining the requirements in question;
2. assess whether the validation programme established (protocol) and the experimental work performed can be considered sufficient to document compliance with the established requirements and needs;
3. assess whether the result of the validation has shown that the established requirements and needs have been met.

6 REVIEW OF PERFORMANCE CHARACTERISTICS FOR A METHOD OF ANALYSIS

The parameters that characterise a method of analysis and which therefore define the suitability of the method of analysis in a given test situation are examined in the following section.

The following performance characteristics are considered:

selectivity, measuring range, linearity, sensitivity, limit of detection, limit of quantitation, robustness, trueness (expressed in the form of bias), precision, measurement uncertainty and accuracy.

The performance validation exercises must encompass all stages of preparation, extraction and analysis encountered by samples during routine analyses.

6.1 Selectivity

6.1.1 Definition

See point 4.26 (including the reference to the term “specificity”)

6.1.2 Content

The validation programme for selectivity/specificity should comprise investigation of possible interference. This should as a minimum encompass:

- investigation of the pure (respective) matrix or a natural specimen with the lowest possible known content;
- investigation of a solution of the relevant analytical parameters in a simple medium;
- investigation of a matrix spiked with pure analytical parameters.

COMMENT: Potential interference may be assessed by spiking the sample matrix with the interfering substances under consideration.

6.1.3 Scale

- In the case of a recognised method of analysis (cf. section 5.3), in which specific interference problems are mentioned, and where there is no justified suspicion of additional interference, the laboratory is not required to carry out further interference testing.
- In cases involving the use of a method of analysis developed by the laboratory itself, foreseeable interference problems must be investigated and documented, (6.1.2).
- If a specific usage situation might give rise to suspicions of additional interference problems, these must be clarified by supplementary validation before the method of analysis is adopted.
- The laboratory's declaration concerning the scale of validation chosen for the method of analysis in question must incorporate considerations of possible interference – e.g. chemical, spectral, chromatographic, detection-related, extraction-related, distillation-related, etc. Interference is assessed with reference to (where possible) specifically described test matrices, and (where relevant) with low concentrations of the analytical parameter(s) concerned.

6.2 Measuring Range

6.2.1 Definition

See point 4.15

6.2.2 Content

To establish the measuring range, trueness and precision are determined with specimens whose concentrations represent the desired measuring range. The measuring range of a method of analysis includes, where appropriate, dilution steps.

6.2.3 Scale

The laboratory must have established a validation programme (protocol) with details of how validation of the measuring range of the method of analysis can be established. This includes documentation of linearity, trueness and precision [see 6.3, 6.8, 6.9].

Trueness and precision must as a minimum be documented at the lowest and highest concentration level of the measuring range. It will often be relevant for the laboratory also to have undertaken supplementary determinations at a number of levels within the measuring range.

If the laboratory uses the method of analysis on matrices whose content of analytical parameters are beyond the measuring range, and which the laboratory consequently wishes to dilute, additional validation/verification must be undertaken.

6.3 Linearity/non-linearity

6.3.1 Definition

See points 4.13

6.3.2 Content

Evaluation of linearity can be performed in various steps of a method of analysis: linearity can be investigated for calibration function and thus only cover instrumental measurement, or linearity can be investigated for the method as a whole and thus be an investigation of trueness as a function of the concentration of the specimen. The investigations described here are aimed at establishing the shape of the calibration function. The principles can, however, to a large extent also be applied in assessing linearity for the method as a whole.

The graphical shape of calibration curves (linearity/non-linearity) is established by analysing a number of calibration standards that cover the entire desired instrumental measuring range. Calibration standards may be reference materials containing the analytical parameters in question in the correct concentration range in relevant matrices. ‘In-house’ calibration standards at the desired concentration level are generally used, preferably consisting of the correct matrix.

In the event of ‘pure’ solutions being used as calibration standards, there is a risk of matrix effects that are not taken into account. In this event, particular care must be taken in determining the trueness of the method.

6.3.3 Scale

The calibration function forms the basis for all subsequent quantitative determinations, and the laboratory must therefore have displayed particular care in validating the process. Experience shows that ignorance of the precise shape of the calibration curve is a source of erroneous results in subsequent routine use of the method of analysis.

It is important that the linearity investigation is performed with calibration standards of the same type as those used in routine analysis.

There are many methods for calibrating and assessing linearity, partly dictated by the various analytical techniques and their specific use. This document does not therefore set out generally applicable

guidelines appropriate to all cases. All laboratories must, however, have undertaken detailed consideration of how the validity of the calibration function adopted and its graphical course are investigated and documented. As a minimum, such considerations should be based on the following approach:

1. **Measurement of a minimum of 6 calibration standards** evenly distributed over the measuring range. The measurements should as far as possible be performed in one series with at least double determinations (where, appropriate triple determinations) at each concentration level and in random order. If the method of analysis is to be used at low concentrations, the measurements should further comprise supplementary concentration levels in this range.
2. **Graphical representation of response as a function of concentration** – both for individual values and for the mean.
3. **Assessment of** whether there is (reasonable) **variance homogeneity** throughout the measuring range (based on the spread at individual points).
4. **Performance of linear regression.** The regression calculation should be weighted if the assessment of variance homogeneity under point 3 indicates its relevance.
5. **Plotting of residual values** relative to the regression line.
6. **Assessment of any outliers.** Following possible rejection or replacement of one or more data points, parts 4 – 6 above shall be repeated. (Note that rejection should normally only have taken place if the laboratory has been able to identify a cause for the extreme value, and if it otherwise leaves sufficient points for performing this procedure).
7. **Assessment of linearity** (based on a plot of the mean, see part 2 above) and residual plot).
8. If a linear link cannot be substantiated, the laboratory may choose one of several approaches:
 - Reduction of the measuring range used and subsequent documentation of linearity;
 - Description of a piecemeal linear link, with the difference in gradient coefficients being assessed;
 - Determination of a second- or third-level polynomial function and subsequent assessment of residuals (NB: normally requires a minimum of 9 or 12 data points respectively; fourth-level polynomials and above will normally be unrealistic).

Since, according to experience, many laboratories assess the level of linear correlation exclusively by examining the correlation (r) or determination coefficient (r^2) for the data set in question, it should be emphasised that this does not provide adequate documentation of linearity. The laboratory shall, as a minimum, also have undertaken a visual assessment of plots, but in addition r or r^2 can certainly be used as a supplementary parameter in the assessment. A ‘goodness-of-fit’ or lack-of-fit test may be utilised to underpin the visual assessment.

The calibration standards used by the laboratory must be of the same type, which are subsequently used in routine analysis.

For certain types of method, the laboratory shall have assessed the calibration function of the method by studying a graphical representation of response factors at various levels. The term “response factor” should in this context be understood to mean the relationship between response for a given analytical parameter and the actual content of the analytical parameter (e.g. the content in certified reference material or spiked content in natural specimens).

If the laboratory has used, for example, instrument-integrated software for automatically calculating regression, it should at the very least have been ascertained that various plots are suitable for a visual assessment, and that algorithms used are otherwise in accordance with the above principles.

The investigations performed document the waveform of the calibration function, but not necessarily that the calibration leads to a correct result. This will only become apparent when the trueness of the method is investigated (see section 6.8).

If the laboratory has, on the basis of the above, documented and described a linear relationship within the measuring range, subsequent routine calibration of equipment may often be simpler (e.g. 2-3 point calibration incorporating the lowest and highest calibration points which would be used in the full calibration), with the laboratory nevertheless having taken account of any lack of variance homogeneity in the measuring range. Investigations to validate the method shall accordingly be performed with the calibration procedure chosen for routine use.

6.4 Sensitivity

6.4.1 Definition

See point 4.7

6.4.2 Content

Sensitivity is calculated with the aid of the gradient of the calibration curve with determination of the linear range.

Sensitivity will be dependent on the equipment and its optimisation.

6.4.3 Scale

In so far as sensitivity affects one or more of the other performance characteristics, it shall be laid down with the associated acceptance limit. It shall be documented to what extent sensitivity affects other performance characteristics.

6.5 Detection Limit

6.5.1 Definition

See point 4.5

The laboratory shall follow the definition that generally applies to the testing field in question.

6.5.2 Content

The detection limit is based primarily on the standard deviation for natural specimens without – or with a very low – content of the analytical parameter. Alternatively, blank specimens or natural specimens or homogeneous inspection materials with a very low content of the analytical parameter may be used, with the term ‘low content’ being understood to mean a content that is as close as possible to the expected detection limit. These may, however, give lower values for the detection limit.

In determining the detection limit, every specimen shall have undergone the complete method of analysis.

If a definition of detection limit based on the desire to avoid false positives or false negatives is used, the standard deviation used shall be based on the repeatability of the method.

If a definition of detection limit based on the desire to limit the uncertainty of the measurement is used, the standard deviation should be based on the laboratory reproducibility of the method, i.e. a standard deviation that caters both for variation within a day and day-to-day variation. Within certain testing fields, there is, however, a tradition or even a requirement to use repeatability standard deviation instead of total standard deviation. The choice of measure for variation shall therefore be determined in accordance with what is the current standard or requirement for the testing field.

The laboratory shall substantiate the choice of definition of detection limit and the choice of measure for variation in its calculation.

Determination of standard deviation is carried out in the same way as described in section 6.9, Precision.

- If the detection limit (DL) is based on the desire to avoid false positives or false negatives, calculation should be performed as follows:

$$DL = 2 \cdot t_{1-\alpha}(f) \cdot s_r$$

Where:

- 2 is due to the desire to have a risk of false negatives of α . This (α) is also the risk of false positives at a concentration equal to $1 \cdot t_{1-\alpha}(f) \cdot s_r$. At a concentration equal to the detection limit, the risk of false positives is at least 10 times less than α .

$t_{1-\alpha}$ the tabulated value in the Student's t-distribution with accepted risk α for false negatives and f degrees of freedom

f the number of degrees of freedom for determining s_r ,

s_r the repeatability standard deviation for a natural specimen without a content or a specimen with a content of the analytical parameter as close as possible to the expected detection limit

- If a blank value is deducted in the calculation of the analytical result, the following expression should be used for calculation purposes:

$$DL = 2 \cdot t_{1-\alpha}(f) \cdot \sqrt{s_r^2 + s_{r,\text{blank}}^2} \cdot \sqrt{1 + 1/n},$$

Where:

$s_{r,\text{blank}}$ the repeatability standard deviation for a blank specimen, and

n the number of blank specimens in a series in routine analysis.

- If the detection limit (DL) is based on the desire to limit the uncertainty of the measurement, the calculation should be performed as follows:

$$DL = \text{blank value} + k.s$$

Where:

k a factor that is multiplied by the standard deviation to calculate the uncertainty. In this context, 3 is traditionally used.

s standard deviation for a natural specimen without content, for a specimen with a very low content of the analytical parameter or for blank specimens. The standard deviation should be the laboratory reproducibility unless the laboratory with technical justification has argued for use of the repeatability standard deviation.

In most cases, the blank value is deducted from the measurement result before final calculation of the analytical result. If this is the case, it can also be deducted from the detection limit, which thereby only becomes $k.s$.

The laboratory should be alert to sector-specific requirements for defining the detection limit.

6.5.3 Scale

The laboratory shall have validated/verified the detection limit in all cases where a method of analysis is used at low concentrations, with the term 'low concentrations' in this context being understood to mean concentrations less than 5 times the expected detection limit.

The standard deviation, which is used as the basis for calculating a detection limit, shall be determined with at least 6 degrees of freedom. This may, for example, have occurred through the performance of 7 determinations in a single analytical series or by double determinations in each of 6 analytical series.

6.6 Limit of Quantitation

6.6.1 Definition

See point 4.10

6.6.2 Content

The limit of quantitation should as far as possible be based on laboratory reproducibility for natural specimens without content of the analytical parameter, blank specimens or natural specimens or homogeneous inspection materials with a very low content of the analytical parameter, i.e. the same specimens as used in determining the detection limit. The standard deviation is determined as described in section 6.9 Precision.

The quantitation limit is determined traditionally as the blank value plus 10 times the repeatability standard deviation, or in some contexts 3 times the detection limit is used, which gives largely the same figure. Other factors can be used within certain testing fields and the laboratory shall in that case refer to the factor that is current within the testing field in question. In many cases, the blank value is deducted from the measurement result before final calculation of the analytical result. If this is the case, it can also be deducted from the quantitation limit, which thereby only becomes a factor times s.

6.6.3 Scale

Many testing fields do not operate with the limit of quantitation. The limit of quantitation shall therefore only be included in validation if this is the requirement for the testing field in question.

6.7 Robustness

6.7.1 Definition

See point 4.25

6.7.2 Content

Changes in the experimental conditions for a method of analysis, such as source and age of reagents, pH, specimen composition and temperature, etc. may possibly affect the results of the method, as may occur in the performance of the method of analysis by other laboratories.

It is possible to systematically investigate whether small changes in the procedure result in significant changes in the analytical result in a robustness test.

The effects on the analytical method are investigated in this way when changes are made in either surroundings and/or in the method procedure – relative to precision and/or accuracy.

For a more detailed description of a test set-up for robustness testing, see the "Statistical Manual of the AOAC" [9]

6.7.3 Scale

The conditions that may have a significant effect on the results of the method of analysis shall be identified.

Trials shall be conducted with the aid of blank specimens, (certified) reference material, specimens of known composition, etc., in which one or more changes in the procedure have been made and the effect on the analytical result has been assessed. In the case of a recognised method of analysis for which specific robustness problems are mentioned, and for which there is no justified suspicion of other problems, the robustness test is considered adequate.

6.8 Trueness/Bias

6.8.1 Definition

See points 4.3 and 4.9

6.8.2 Content

The description of the trueness of the method of analysis can be visualised by determining the bias of the method by at least 2 of the following procedures:

- use of a certified reference material;
- comparison of the method with a primary method or a reference method with little or no bias;
- use of traceable, commercially available reference materials or internal inspection materials with an assigned value;
- use of spiked specimens, based on blank or positive specimens;
- participation in a performance testing programme in which setting of the reference values meets the quality criteria specified in ISO Guide 43 [13,14].

If possible, several of the above-mentioned procedures shall be incorporated in the determination of the bias of the method of analysis.

6.8.3 Scale

To take account of any variation between runs, bias shall be determined over several days and preferably throughout the measuring range through the use of a suitable combination of different specimens. As the method of analysis cannot be expected to have the same bias throughout the measuring range (e.g. with non-linear calibration curves), several concentration levels shall be incorporated in the determination of bias (at least one determination at a high and a low level). Otherwise, the laboratory should be able to prove that the method of analysis employed has the same trueness throughout the measuring range.

Bias can be calculated as the relative deviation from a reference value, measured as a mean determined with at least 6 degrees of freedom. Bias shall be tested with a t-test or by assessment of the observed value's confidence interval (calculated from the laboratory reproducibility standard deviation) relative to the reference value and its indication of uncertainty.

If the laboratory has used spiking trials in connection with documentation of the trueness of the method of analysis, the laboratory shall be able to prove that certain relevant considerations have been catered for concerning the problems associated with spiking and recovery of pure analytical parameters in relation to the presence of naturally occurring analytical parameters.

6.9 Precision

6.9.1. Definition

See points 4.18

6.9.2 Content

Repeatability:

The repeatability standard deviation, s_r , is determined as the standard deviation for a suitable number of determinations measured under repeatability conditions.

Laboratory Reproducibility:

The reproducibility standard deviation, s_R , is determined as the standard deviation of a suitable number of determinations on identical specimens analysed over several days with at least 2 different calibration standards. In addition, the analyses should as far as possible be performed by several competent

individuals. If the method of analysis is to be used on more than one item of equipment (e.g. several GCs), the relevant measuring equipment must be used for determining s_R .

6.9.3 Scale

The repeatability standard deviation shall be determined with at least 6 degrees of freedom. This can be achieved, for example, by analysing 7 times in series with one specimen ($f=6$), 4 times in series with 2 specimens ($f=6$), 3 times in series with 3 specimens ($f=6$), etc.

In the case of s_r , it is therefore the case that this can be determined as the standard deviation of several specimens run in the series or as the pooled standard deviation of a number of multiple double determinations run over several series.

The number of degrees of freedom achieved with various combinations of numbers of series and numbers of specimens in each series is set out in the diagram below:

Number of specimens per series	Number of series	Degrees of freedom for repeatability standard deviation	Degrees of freedom for laboratory reproducibility standard deviation
7	1	6	Not determined
4	2	6	7
3	3	6	8
2	6	6	11
N	M	(n-1)*m	n*m - 1

As repeatability and reproducibility will typically vary within the measuring range of a method of analysis, these should be determined at several concentration levels (min. 3), with one of these levels being close to the lower value for the measuring range.

With variations within the matrix in question, determinations at several levels will be relevant.

All the matrices for the method of analysis (or groups thereof) should form part of the laboratory's considerations with respect to validating the precision of the method.

6.10 Measurement Uncertainty

This subject is not dealt with in detail in this guideline, but documentation of measurement uncertainty should form an integral part of method validation. It will be natural for the laboratory to combine a method validation study with the formulation of an uncertainty budget, not least owing to the fact that data from the experimental work will form the basis for formulating an uncertainty budget.

6.10.1 Definition

See point 4.16

6.10.2 Content

As is apparent from the definition (4.16), measurement uncertainty is a property of a measurement result. A method of analysis cannot therefore be assigned measurement uncertainty, but forms a source of uncertainty that is usually characterised by its repeatability/reproducibility (ISO 5725-1 [10]). The contribution of the method of analysis to the uncertainty of the analytical result depends on many other things, as is apparent from the guidance in the GUM document [15] and the EURACHEM/CITAC document [16].

Determination of measurement uncertainty comprises the following steps:

1. Identification of all sources of uncertainty;
2. Quantitation of uncertainty components;

3. Calculation of the combined measurement uncertainty for the method of analysis.

The procedure for determining measurement uncertainty in connection with chemical methods of analysis should as far as possible follow the guidelines in the EURACHEM/CITAC guide [16].

6.10.3 Scale

Following identification of all sources of uncertainty, it shall be assessed whether individual components occur several times, whether the contribution from several components could, where appropriate, be grouped and whether uncertainty components and groups of them can be quantified on the basis of data already available – including data from method validation. For example, determination of laboratory reproducibility may cover uncertainty contributions from sub-specimen collection, volumetric equipment and weights used in specimen preparation (but not in preparation of the calibration standards), contributions from random noise in the detection process, contamination/loss of analytical parameters, recovery, estimation of calibration curve and blank correction.

After the determination of measurement uncertainty has been simplified, the components left shall be quantified. Lastly, the model according to which the contributions are combined shall be taken into account.

When reporting measurement uncertainty, so-called “expanding measurement uncertainty” shall be used. This is calculated by multiplying the combined standard uncertainty with a coverage factor, k , normally 2.

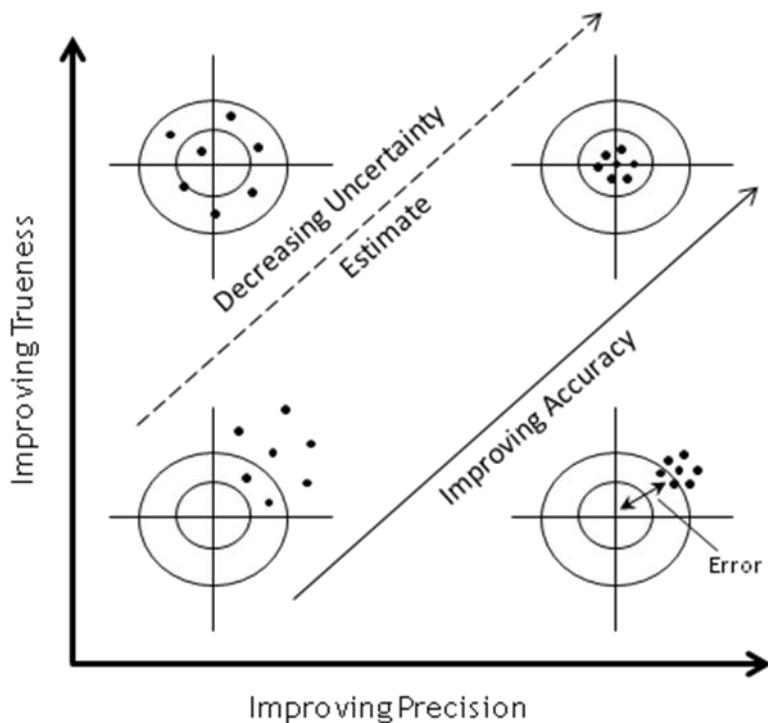
6.11 Accuracy

Accuracy is (like measurement uncertainty) a property of a measurement result, and is therefore not included as such in the method validation itself.

6.11.1 Definition

See point 4.17

The link between the terms *precision*, *trueness* and *accuracy* can, among other things, be illustrated by the following figure:



6.11.2 Content

Accuracy can be described on the basis of the bias of the method of analysis (trueness) and precision.

6.11.3 Scale

The determination of accuracy does not require extra validation work.

7. FOLLOW-UP METHOD VALIDATION

The laboratory shall continually check that a method of analysis meets the values for the performance characteristics documented in connection with validation.

The laboratory shall have established and implemented a procedure for internal monitoring in relation to relevant performance characteristics for the method. In this context, it should be stressed that calibration standards and control specimens must have a different origin.

8. LABORATORY COMPETENCE

The laboratory shall have appointed individuals (individuals responsible for the method) who can assume responsibility for performing validation work on the basis of documented knowledge and experience, including:

- technical (theoretical and practical) knowledge of the area of analysis in question and analytical parameters;
- an understanding of the intended use of the method of analysis – e.g. via contact with customers;
- knowledge of the validity of methods of analysis.

A declaration concerning the validity of a method of analysis is purely the responsibility of the individual responsible for the method, and this individual also has overall responsibility for the entire course of

validation. The practical work on performing the validation work itself can be delegated to other competent staff within the laboratory.

The individual responsible for the method may, where appropriate, be the same party who, for accreditation purposes, authorises the method of analysis.

9. LABORATORY DOCUMENTATION

9.1 Procedure(s)

The laboratory shall have one or more procedure(s) defining the laboratory's approach in connection with the performance of validation work. Such procedures shall as a minimum contain the laboratory's guidelines for:

- appointment of an individual responsible for the method validation work (individual responsible for the method);
- specification of the customer's/laboratory's requirements concerning the testing in question;
- guidelines for laying down the scale of validation with a view to determining performance characteristics of the method of analysis in relation to the requirements in question;
- drafting of a protocol describing the validation work;
- investigation of whether the customer's/laboratory's requirements can be met with the method of analysis after the performance requirements have been laid down;
- drafting a declaration concerning the validity and approval of the method of analysis.

Different validation procedures may be necessary / appropriate for different testing fields.

9.2 Reporting

The laboratory shall, for every validation job, have drawn up a validation report, which shall as a minimum contain:

- a declaration concerning the scale of validation decided upon;
- a protocol for the validation work in question (or a reference thereto);
- results of validation work (or reference thereto);
- a final assessment of these results;
- a declaration concerning the validity of the method of analysis with, where appropriate, release of the method (approval).

The report shall be signed by the individual responsible for the method (cf. point 8) and must be kept in the laboratory as part of the documentation for the method of analysis in question.

10. Contact

For further information about this statement please contact an INAB officer at **The Irish National Accreditation Board**.

Phone: 1890 289 389
E-mail: inab@inab.ie
Website: www.inab.ie