GUIDE 33

Third edition
2015-02-01

Reference materials — Good practice in using reference materials

Matériaux de référence — Bonne pratique d'utilisation des matériaux de référence

Reference number
ISO GUIDE 33:2015(E)

International Organization for Standardization

COPYRIGHT PROTECTED DOCUMENT

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword ..... **v**

Introduction ..... **vi**

1 Scope **1**

2 Normative references **1**

3 Terms and definitions **2**

4 Symbols **3**

5 Conventions **3**

6 RMs and their role in measurement **4**
6.1 Common applications of RMs 4
6.2 Property values 6
6.2.1 General 6
6.2.2 Specification of the property 6
6.3 Uncertainty statement 6
6.4 Traceability statement 8

7 Handling of RMs and CRMs **9**

8 Assessment of precision **9**
8.1 General 9
8.2 Number of replicate measurements 9
8.3 Requirements with respect to the RM 10
8.4 Measurement 11
8.5 Data treatment 11
8.6 Calculation and assessment of precision 12

9 Bias assessment **12**
9.1 General 12
9.2 Approach to bias checking 13
9.3 Utilizing bias data 13

10 Calibration **14**
10.1 General 14
10.2 Establishing metrological traceability 15
10.3 Calibration models 15

11 Assigning values to other materials **15**
11.1 General 15
11.2 Pure materials 16
11.3 Gravimetry and volumetry 17

12 Conventional scales **18**
12.1 General 18
12.2 pH-scale 19
12.3 Octane number 19

13 Selection of CRMs and RMs **20**
13.1 General 20
13.2 Selection of a CRM 20
13.3 Selection of RMs 23
13.4 Relevance to the measurement system 23

Annex A (informative) **Key characteristics of a reference material with respect its common applications** **25**

Annex B (informative) **Calibration models and associated uncertainty models** **26**

Annex C (informative) **Decision errors** **29**

Bibliography **30**

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 on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/REMCO, *Committee on reference materials*.

This third edition cancels and replaces the second edition (ISO Guide 33:2000), and ISO Guide 32:1997 which have been technically revised.

Introduction

The aim of this Guide is to provide general recommendations on the use of RMs. These recommendations are exemplified by real-world examples, which to some degree also reflect the level of complexity associated with RMs. This level of detail is deemed to be useful for anyone who has a responsibility in the quality management in laboratories, such as drafters, reviewers, managers, and assessors of procedures, working instructions, standard operating procedures and the like.

The main applications of reference materials are calibration, establishing traceability, method validation, assigning values to other materials, and quality control.

Reference materials — Good practice in using reference materials

1 Scope

1.1 This Guide describes good practice in using reference materials (RMs), and certified reference materials (CRMs) in particular, in measurement processes. These uses include the assessment of precision and trueness of measurement methods, quality control, assigning values to materials, calibration, and the establishment of conventional scales. This Guide also relates key characteristics of various types of RMs to the different applications.

1.2 For CRMs, the metrological traceability of the property values to international scales or other measurement standards has been established. For RMs not being CRMs, this kind of traceability of property values has often not been established. Nevertheless, these RMs can still be used for assessing parts of measurement procedures, including evaluating various levels of precision

1.3 Mainstream applications of RM include precision control (Clause 8), bias assessment (Clause 9), calibration (Clause 10), preparation of calibration RMs (Clause 11) and maintaining conventional scales (Clause 12).

NOTE Not all types of RMs can be used for all indicated purposes.

1.4 The preparation of RMs for calibration is also part of the scope of ISO Guides 34[1] and 35[2]. The treatment in this Guide is limited to the fundamentals of small-scale preparation of RMs and the value assignment, as used by laboratories to calibrate their equipment. Larger scale production of such RMs, with the possible aim of distribution is beyond the scope of this Guide. This type of activity is covered in ISO Guides 34[1] and 35[2].

1.5 The development of working standards, as used in, e.g. natural gas analysis, clinical chemistry, and the pharmaceutical industry is not covered in this Guide. This type of activity is covered in ISO Guides 34[1] and 35[2].

2 Normative references

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO Guide 30, *Terms and definitions used in connection with reference materials*

ISO/IEC Guide 98-3, *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*

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

NOTE The “Guide to the expression of uncertainty in measurement” is referred to as “GUM”, whereas the “International vocabulary of basic and general terms in metrology” is referred to as “VIM”.

3 Terms and definitions

For the purposes of this document, the terms and definitions in ISO/IEC Guide 98-3, ISO/IEC Guide 99 and ISO Guide 30 and the following apply.

NOTE Further definitions can be found in the ISO online browsing platform, accessible through https://www.iso.org/OBP/ui/

3.1
reference material
RM
material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process

Note 1 to entry: RM is a generic term.

Note 2 to entry: Properties can be quantitative or qualitative, e.g. identity of substances or species.

Note 3 to entry: Uses may include the calibration of a measurement system, assessment of a measurement procedure, assigning values to other materials, and quality control.

Note 4 to entry: ISO/IEC Guide 99:2007, has an analogous definition (5.13), but restricts the term "measurement" to apply to quantitative values. However, Note 3 of ISO/IEC Guide 99:2007, 5.13 (VIM), specifically includes qualitative properties, called "nominal properties".

[SOURCE: ISO Guide 30]

3.2
certified reference material
CRM
reference material characterized by a metrologically valid procedure for one or more specified properties, accompanied by an RM certificate that provides the value of the specified property, its associated uncertainty, and a statement of metrological traceability

Note 1 to entry: The concept of value includes a nominal property or qualitative attribute such as identity or sequence. Uncertainties for such attributes may be expressed as probabilities or levels of confidence

Note 2 to entry: Metrologically valid procedures for the production and certification of RMs are given in, among others, ISO Guides 34 and 35.

Note 3 to entry: ISO Guide 31[17] gives guidance on the contents of RM certificates.

Note 4 to entry: ISO/IEC Guide 99:2007 has an analogous definition (5.14).

[SOURCE: ISO Guide 30]

3.3
property value
<of a reference material> value corresponding to a quantity representing a physical, chemical or biological property of a reference material

[SOURCE: ISO Guide 30]

3.4
certified value
value, assigned to a property of a reference material (RM) that is accompanied by an uncertainty statement and a statement of metrological traceability, identified as such in the RM certificate

[SOURCE: ISO Guide 30]

3.5
indicative value
information value
informative value
value of a quantity or property of a reference material, which is provided for information only

Note 1 to entry: An indicative value cannot be used as a reference in a traceability chain

[SOURCE: ISO Guide 30]

3.6
calibrant
reference material used for calibration of equipment or a measurement procedure

[SOURCE: ISO Guide 30]

2.7
quality control material
reference material used for quality control of a measurement

[SOURCE: ISO Guide 30]

4 Symbols

α	risk of error of the first type (type I error)
β	risk of error of the second type (type II error)
χ^2	Chi-squared
d	measurement bias
k	coverage factor
s_{w}	standard deviation computed from repeat observations
σ_{wo}	required within-laboratory standard deviation
$u()$	standard uncertainty of the parameter in parenthesis
$U()$	expanded uncertainty of the parameter in parenthesis
u_{CRM}	standard uncertainty associated with property value of the CRM
u_{meas}	standard uncertainty associated with value obtained by measuring the CRM
u_{prep}	uncertainty associated with the value obtained from preparation of a calibrant
x_{CRM}	value of a specified property of the CRM
x_{meas}	value obtained by measuring the CRM
x_{prep}	value obtained from preparation of a calibrant
$\overline{x}$	average of repeat observations

5 Conventions

In this Guide, the following conventions are used.

5.1 A measurand is specified in such a way that there exists a unique, but unknowable, 'true value'.

5.2 All statistical methods used in this Guide are based on the following assumptions.

a) The certified value is the best estimate of the true value of the property of the CRM.

b) All variation, be it associated with the material (i.e. homogeneity) or the measurement process, is random and follows a normal probability distribution. The values of probabilities stated in this Guide assume normality. The probability may be different if there is deviation from normality.

5.3 The concept of "certified reference material" (CRM) as used in this Guide also includes RMs with property values that are accompanied by the statements of metrological traceability or measurement uncertainty. These property values are assumed to be obtained through characterization as described in ISO Guides 34[1] and 35[2].

5.4 Where the term RM is used in this Guide, it means that any RM can be used for the purpose described. The use of a CRM is an option, but usually not the most economical one. In practice, in most cases an RM will be used that comes without property values, uncertainties and a traceability statement.

5.5 Values, given as "indicative", "informative", "for information" or otherwise identified as not being covered by the statements of metrological traceability or measurement uncertainty, are considered to be unfit for use in metrological applications requiring a value assigned to the measurand, such as, calibration, or the assignment of values to other materials. These values are however useful to verify whether an RM is suitable for precision control, or other applications that do not require a property value.

5.6 Throughout this Guide, the law of propagation of uncertainty is used. Other methods of propagating uncertainties can be applied as well, and in some cases such alternative methods are required by the circumstances of the application. Further guidance on these matters is given in the GUM and its supplements.

6 RMs and their role in measurement

6.1 Common applications of RMs

6.1.1 RMs, and CRMs in particular, are widely used for the following purposes:

— Calibration of equipment or a measurement procedure (Clause 10);

— Establishing metrological traceability (Clauses 9, 10 and 11);

— Method validation (Clauses 8 and 9);

— Assigning values to other materials (Clause 11);

— Quality control of a measurement or measurement procedure (Clause 8 and 9);

— Maintaining conventional scales (Clause 12).

Figure 1 gives an outline of a measurement, including sampling and sample preparation. The possible role(s) of CRMs are indicated.

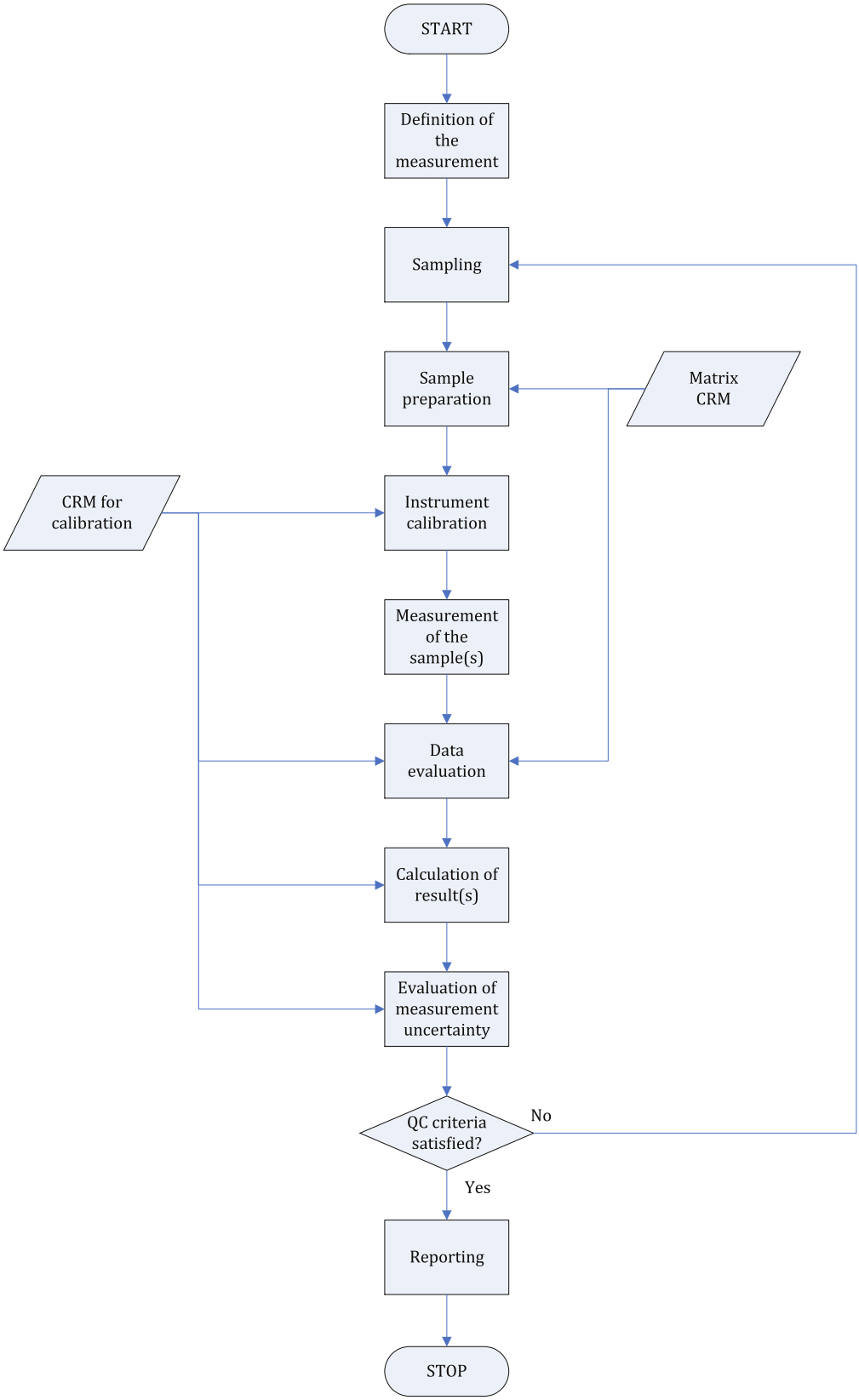

Figure 1 — Schematic outline of a measurement and the role of CRMs therein

6.1.2 Written standards for quality systems of laboratories, e.g. ISO/IEC 17025[4] and ISO 15189[5], require measurement results to be metrologically traceable and measurement equipment calibrated. Metrological traceability is a prerequisite for achieving comparable and compatible measurement results.

EXAMPLE A wine with a volume fraction of 12 % alcohol can be usefully compared with another wine with a volume fraction of 13,5 % alcohol.

6.1.3 The general public takes it for granted that measurement results, expressed in appropriate units, are comparable. In order to meet this implicit expectation of measurement data, laboratories should ensure that all equipment is properly calibrated, using measurement standards which in turn have been made metrologically traceable to the realization of the relevant unit. In many cases, this unit is part of the SI, the International System of units.

6.1.4 A summary list of key characteristics of RMs, cross-referenced to the common applications of RMs, is given in Annex A of this Guide.

6.2 Property values

6.2.1 General

6.2.1.1 CRMs are characterized for one or more properties. These property values are accompanied by

a) a clear specification of the property concerned,

b) an uncertainty statement,

c) a statement of metrological traceability,

d) a period of validity of the certificate.

The user should verify that all this information is available in an unambiguous form.

6.2.1.2 Indicative values should not be used for any of the uses of CRMs described in this Guide.

NOTE The terminology used in practice for indicative values is not always consistent with this Guide.

6.2.2 Specification of the property

6.2.2.1 The unambiguous specification of the property concerned assists greatly in deciding whether the CRM is appropriate for the intended application. The user of a CRM is responsible for assessing the suitability of the material for the intended purpose.

EXAMPLE For trace elements in soil, it is important to specify whether it is the total content, content obtained by incomplete destruction (e.g. *aqua regia*), leachable content, or a particular species containing the trace element.

6.2.2.2 The property values should be given in appropriate units, preferably SI units. The property values should be given in an appropriate number of digits, avoiding unnecessarily loss of accuracy on one hand and giving a false impression of accuracy on the other.

NOTE The GUM (ISO/IEC Guide 98-3:2008, Clause 7) gives guidance on the rounding of measurement results and associated uncertainties.

6.3 Uncertainty statement

6.3.1 The uncertainty statement should be readily understood, which among other considerations requires that all information necessary to convert the uncertainty stated into a standard uncertainty is

available. If an expanded uncertainty is given, then usually the appropriate coverage factor is sufficient for this conversion.

EXAMPLE A calibration certificate for a gas mixture states for the amount of substance fraction of carbon monoxide the following:

x_{CO} = (41 122 ± 28) µmol/mol (k = 2)

The expanded uncertainty is 28 µmol/mol. The standard uncertainty is obtained using

$$u = \frac{U}{k} = \frac{28}{2}\ \mu\text{mol mol}^{-1} = 14\ \mu\text{mol mol}^{-1}$$

6.3.2 If a coverage interval is given, then the (assumed) probability density function of the property value should be specified, including the coverage probability (e.g. 95 %) of the stated interval. Such an interval can be asymmetric. Sometimes, it can be necessary to make additional assumptions concerning, for example, the appropriate coverage factor. In those cases, the guidance of the GUM (ISO/IEC Guide 98-3:2008, 6.3) should be followed.

EXAMPLE The carbon content in a gas coal is specified to be 760,1 mg/g, and the uncertainty is stated to be 2,1 mg/g. The following footnote is added to the uncertainty statement: "The uncertainty is expressed as the 95 % confidence interval. It is applicable when the reference material is used for calibration purposes."

From the certification report, it is clear that the certified values have been obtained from an interlaboratory experiment, and therefore it is reasonable to assume the normal distribution. The half-width of a 95 % confidence interval equals 1,96 times the standard deviation. There is however no practical difference with using a (coverage) factor of 2, which would correspond for the normal distribution with 95,45 % level of confidence.

The standard uncertainty is obtained through $u = \frac{U}{k} = \frac{2,1}{2}\ \text{mg/g} = 1,05\ \text{mg/g}$

6.3.3 The uncertainties stated should be given in the same unit as the property values, or alternatively, expressed as a fraction of the property value (i.e. as a relative expanded uncertainty). If such fractions are used, it should be established whether the conversion into absolute standard uncertainties can be performed in an unambiguous way.

NOTE Such fractions include percentages, per mil, and parts-per-million (ppm), although none of these are recommended for this purpose, because of their ambiguity.

6.3.4 CRMs are accompanied by a certificate stating, among others, the properties certified, their values, and their associated uncertainties (see 6.2.1.1). It is beyond the scope of this Guide to describe how the uncertainty associated with property values is established, but it is important to understand the potential main contributors.

NOTE Details concerning the establishment of an uncertainty budget for property values are given in ISO Guide 35 [2].

The main contributors to the uncertainty associated with the property values of a CRM include:[2]

— uncertainty from characterization;

— uncertainty due to long-term stability;

— uncertainty due to short-term stability (stability of the material under transport conditions);

— uncertainty due to between-bottle variation.

6.3.5 In some cases, detailed knowledge of different parts of the uncertainty budget can be helpful, in particular if such a contribution is the largest. This information may be available from the CRM producer.

6.3.6 Not all uncertainty statements accompanying the property values on certificates of CRMs produced before the late 1990s include effects of batch inhomogeneity and instability. The uncertainty stated on the certificate should comprise all factors that may impact the scatter of the property value(s) across the batch and over time. Ultimately, the uncertainty stated should be applicable to the single package to be used in the measurement process.[2]

NOTE If the uncertainty stated is too small, the expanded uncertainty has a lower level of coverage than stated.

6.3.7 RMs that do not come with property values, should come with some information regarding the (between-bottle) homogeneity and long-term stability of the properties for which the RM can be used. The user should verify whether this information is provided in a form, that it can be used to assess the suitability of the RM. Such assessment may include the use of the information concerning homogeneity and stability in subsequent uncertainty calculations.

6.4 Traceability statement

6.4.1 Metrological traceability is a property of a measurement result. As the value obtained during the characterization of an RM (the property value) is a measurement result, it can have this attribute too. A key characteristic of property values of CRMs is that their metrological traceability is well established.

6.4.2 The user of a CRM should verify that the property values come with a statement concerning the metrological traceability of these values. This statement should inform the user about the measurement scale to which these values refer, so they can verify whether the CRM is suitable for the intended use.

NOTE In most cases, the measurement scale is the SI unit.

6.4.3 To allow interpretation of statements of metrological traceability, the following particulars are necessary, which should be stated on a certificate or other documentation accompanying a CRM:

a) the specification of the measurand;

b) the unit to which the property value is made traceable;

c) the method used for sample handling/transformation and measurement procedures(s)/technique(s) used in the characterization;

d) the approach to characterization (e.g. single method, two methods, multiple laboratories, etc.).

NOTE Documentation accompanying a CRM can be made available by different means, including websites, E-mail, or publications in the open literature.

6.4.4 The provider of the result of a measurement is responsible for supporting the claim of metrological traceability for that result or value. In the case of CRMs, the RM producer bears this responsibility. It is the user's responsibility to review the appropriateness of the metrological traceability for their intended purpose.

6.4.5 For assessing a claim of metrological traceability, the user may require more information than provided on the certificate. A claim of metrological traceability is typically supported by items mentioned in 6.4.3.

The user of an RM should scrutinize this information and assess the fitness for purpose of a particular RM for the intended use. The user of the RM should check the availability of the information specified in this clause to assist in the assessment. Should essential parts of this information be unavailable, the RM may be unsuitable for an application.

7 Handling of RMs and CRMs

7.1 The instructions for use, as well as those for storage should be followed, as they form part of the conditions under which the property values and associated uncertainties are valid. Improper use of RMs and CRMs can be detrimental to the performance of measurement procedures and should be avoided at all times.

7.2 The expiry date on the certificate should be respected. CRMs should not be used beyond this date.

7.3 In particular for CRMs that allow multiple use, it is important to make sure that the container holding the CRM is properly closed, and it is stored in an appropriate manner. In some cases, repackaging of the remaining material can be necessary. Otherwise, the property values stated may become invalid and the CRM unusable or unreliable. The user should follow the instructions provided by the producer in this respect.

7.4 The minimum subsample size should be respected. Smaller subsamples may be unrepresentative.

7.5 Sub-sampling of such CRMs should be done in a way that the sub sample taken for use reflects the properties of the complete package. Otherwise, over time, the remaining material of the CRM may no longer be representative for the batch that has been produced and certified, and hence, the values and uncertainties stated on the certificate are no longer valid.

NOTE 1 Re–homogenization of CRMs is usually necessary, before sub-sampling can be carried out. Such instructions are typically given in the documentation accompanying the CRM.

NOTE 2 So–called “single–shot” CRMs are designed for use as a single portion. Usually, the unit contains sufficient sample for only one or two measurements. However, where single shot materials are designed for use as a single portion, they should not be subdivided.

8 Assessment of precision

8.1 General

8.1.1 Checking of the precision of a measurement procedure as applied by a laboratory involves comparison of the within-laboratory standard deviation under repeatability conditions (or other specified conditions) and the required value of that standard deviation.

NOTE Measures of precision are the standard deviation under repeatability or reproducibility conditions.

8.1.2 Assessment of precision can be part of the activities a laboratory undertakes when developing or validating a method. Such experiments should ideally be run on RMs covering the scope of the method in terms of matrices (or variations within a matrix) and property value levels.[6] The assessment may also involve multiple laboratories. More guidance on assessing precision in interlaboratory studies is given in ISO 5725 [7]-[12].

8.1.3 Results from periodic checks of a measurement procedure may be recorded on a quality control chart. A range control chart may be used for this purpose.[13]

8.2 Number of replicate measurements

8.2.1 For a reliable assessment of precision within a required confidence interval, the necessary number of replicate measurements can be estimated by a χ^2 test. The number of replicate measurements, n, required depends mainly on the values of α and β and the alternative hypothesis chosen for the assessment of precision. A discussion of the probabilities α and β, associated with the type I and type II risks in statistical hypothesis testing is given in Annex C.

8.2.2 Table 1 shows the relation between the degrees of freedom v (where, in this case, $v = n - 1$) and the ratio of the within-laboratory standard deviation of the measurement process, s_w, and the required value of the within-laboratory standard deviation, σ_{wo}, for various values of β at $\alpha = 0{,}05$.

EXAMPLE For $n = 10$ the probability that the variance of the measurement results will pass the appropriate χ^2-test (see 8.6) at $\alpha = 0{,}05$ is no more than 1 % when the within-laboratory standard deviation, σ_w, of the measurement process is equal to or larger than 2,85 times the required value of σ_{wo}.

Table 1 — Ratio of the standard deviation of the measurement process to the required value for various values of β and degrees of freedom v at α = 0,05

v	$\alpha = 0{,}05$			
	$\beta = 0{,}01$	$\beta = 0{,}05$	$\beta = 0{,}1$	$\beta = 0{,}5$
1	159,5	31,3	15,6	2,73
2	17,3	7,64	5,33	2,08
3	6,25	4,71	3,66	1,82
4	5,65	3,65	2,99	1,68
5	4,47	3,11	2,62	1,59
6	3,80	2,77	2,39	1,53
7	3,37	2,55	2,23	1,49
8	3,07	2,38	2,11	1,45
9	2,85	2,26	2,01	1,42
10	2,67	2,15	1,94	1,40
12	2,43	2,01	1,83	1,36
15	2,19	1,85	1,71	1,32
20	1,95	1,70	1,59	1,27
24	1,83	1,62	1,52	1,25
30	1,71	1,54	1,46	1,22
40	1,59	1,45	1,38	1,19
60	1,45	1,35	1,30	1,15
120	1,30	1,24	1,21	1,11

8.3 Requirements with respect to the RM

8.3.1 For assessing measurement precision, an RM is needed with sufficient homogeneity and stability. The stability of the RM with respect to all properties of interest should at least be sufficient for the period of time that the precision checking measurements take. If necessary, specific precautions should be taken to monitor the stability of the RM used. Such precautions may consist of demonstrating the stability of the measurement process under study by other means, such as the use of a CRM, or by using an other measurement process of which the stability has been demonstrated.

8.3.2 When using an RM for control charting, discrepant results can be caused by stability issues of the RM, rather than problems with the measurement system. Users of RMs should be aware of this possibility and include it in the root cause analysis.

8.3.3 Although the RM used for assessing precision taking should not necessarily have known, metrological traceable property values for the properties of interest, precision measures may be dependent on the nominal value of the measurand, so knowledge about the nominal values of the parameters of interest is usually required to assess the appropriateness of the RM selected for the checking.

8.3.4 More generic guidance on aspects to consider concerning the suitability of the RM can be found in Clause 13.

8.4 Measurement

8.4.1 The user should perform independent replicate measurements. "Independent", in a practical sense, means that a replicate result is not influenced by previous measurements. To perform replicate measurements means to repeat the whole procedure. For example, in the chemical analyses of a solid material, the procedure should be repeated from the weighing of the test portion to the final reading or calculation of the result.

EXAMPLE If the measurement of lead in a soil RM consists of sub-sampling, destruction of the test portions, followed by measuring the aliquots, then the results are independent as far as sub-sampling, destruction, and measurement are concerned. If, in contrast, one aliquot is measured repeatedly, then the standard deviation obtained only covers repeatability effects of measuring an aliquot.

Before starting experimental work, it is very important to check what part of the process is described by the standard deviation σ_{wo}. In many written standards for testing, the repeatability stated refers to the whole standard test method. Consequently, the measurements for assessing the repeatability of such a test method should be carried out accordingly, viz., repeating the test method for each and every (sub) sample completely.

8.4.2 Independent replicate measurements can be achieved in various ways depending on the nature of the process. Parallel replication is not always recommended however, because an error committed at any step of the procedure could affect all replicates. Furthermore, steps like, e.g. the calibration of the measurement equipment, may also need to be included in the replication process.

EXAMPLE In the case of iron ore analyses, replication of the analytical procedure is carried out at different times and includes appropriate calibration. In this case, the standard deviation calculated from the replicates also contains contributions from day-to-day variability and calibration.

8.5 Data treatment

8.5.1 The data thus obtained should first be scrutinised for any irregularities. Data identified as technically invalid should be discarded, regardless of whether or not they seem to fit in the (assumed) probability distribution of the complete dataset. Technically invalid results are usually due to irregularities during the sample processing and/or measurement.

8.5.2 A second type of irregularity includes observations whose values seem to be far away from the other observations in the dataset. Often, but not always, a technical reason can be found and then these data should be discarded. If no technical explanation can be found, the data can be scrutinized for possible outliers using the methods described in ISO 5725-2[8] or ISO 16269-4 [14]. Outliers should be discarded or, in rare cases (e.g. calculation errors), be replaced by corrected data. Whenever possible, outliers should be removed only on the basis of the outcome of more than one outlier test. As a rule, stragglers should be kept in the dataset.

NOTE 1 An excessive number of suspected outliers indicates problems in the measurement process.

NOTE 2 Removing outliers, and removing stragglers in particular, will reduce the scatter in the dataset and may consequently lead to a value of the standard deviation of the measurement process under study that is too small.

NOTE 3 The use of most outlier tests requires an assumption concerning the (expected) shape of the probability density function of the data. If such an assumption is inconsistent with the nature of the data, such an outlier test cannot be applied.

8.6 Calculation and assessment of precision

8.6.1 The precision of the measurement process is assessed by comparing the within-laboratory standard deviation under repeatability conditions with the required value of the within-laboratory standard deviation, σ_{wo}.

Compute the average, $\bar{x}$, and standard deviation, s_w:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

$$s_w = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \tag{2}$$

where

x_i is the individual result;

n is the number of results excluding outliers.

8.6.2 Compute the following ratio:

$$\chi_c^2 = \frac{s_w^2}{\sigma_{w0}^2} \tag{3}$$

where σ_{wo} is the required value of the within-laboratory standard deviation.

$$\chi_{table}^2 = \frac{\chi_{(n-1);0.95}^2}{n-1} \tag{4}$$

denotes the 0,95th quantile of the χ^2 distribution with $(n - 1)$ degrees of freedom, divided by the degrees of freedom $(n - 1)$

Interpretation of χ^2:

$\chi_c^2 \leq \chi_{table}^2$ There is no evidence that the measurement process is not as precise as required.

$\chi_c^2 > \chi_{table}^2$ There is evidence that the measurement process is not as precise as required.

NOTE Values for χ^2 may be taken from tables or computed by software. They can be found at many places including Reference [15].

9 Bias assessment

9.1 General

9.1.1 Checking bias is a key application in laboratories. It can be performed as part of ensuring the quality of measurement results, method validation, or both. For bias checking, it is essential that the reference against which the bias is checked is reliable and metrologically traceable.

9.1.2 The CRM to be used for bias assessment is suitable if the type of material and the properties of interest are relevant for the intended use. The user should confirm the suitability of the CRM prior to the bias assessment.

9.1.3 Quality control materials (QCMs) and other uncharacterised RMs can be used for the assessment of precision (see Clause 8). but due to the lack of a metrologically traceable property value, cannot be used to assess the measurement bias.

NOTE Preparation of quality control materials is covered in ISO Guide 80[3].

9.1.4 In this clause, guidance is given on bias assessment. Determining precision is dealt with in Clause 8 of this Guide.

9.2 Approach to bias checking

9.2.1 The use of a CRM for the purpose of bias checking contributes to the metrological underpinning of the measurement result. It is an essential activity in the validation of the measurement procedure. If measurement results using the same CRM, targeting the same measurand, are metrologically traceable, these results should be free of significant bias. In this case, the measurement procedures generate results that are traceable to the same reference.

9.2.2 The observed difference between the measured value and the property value stated on the certificate should be smaller than the standard uncertainty associated with the difference, i.e.

$$\left|x_{\text{meas}} - x_{\text{CRM}}\right| \leq k\sqrt{u_{\text{meas}}^2 + u_{\text{CRM}}^2} \tag{5}$$

NOTE The choice of a coverage factor requires an (assumed) probability density function and a level of coverage. Often 95 % is used as the coverage level, but this choice may depend on the application.

9.2.3 If the condition (5) holds, then the measured and property values are consistent with one another within their respective uncertainties. As the property value of the CRM is metrologically traceable to some stated reference, ideally the SI, under this condition the result obtained for the CRM confirms the metrological traceability of the results obtained from the measurement procedure.

9.2.4 Users should calibrate their equipment independent of the CRM used for bias checking. The laboratory subsequently assesses the correctness of carrying out all steps in the measurement procedure by comparing the result with the stated property value. If the laboratory can use the same measurement procedure for its routine samples, it can demonstrate metrological traceability of its result to the stated reference for the property value of the CRM.

NOTE 1 In practice, it may be necessary to modify some parts of the measurement procedure when applying it to the CRM. Modifications in the routine measurement procedure necessary for measuring the CRM may compromise the validity of the assessment. The more (or severe) modifications are necessary, the less useful the assessment becomes.

NOTE 2 In cases when the measurement procedure defines the measurand, e.g. enzyme measurements, changes in the measurement procedure will change the definition of the measurand.

9.2.5 The evaluation of method bias as outlined in this clause is not limited to the use of a single CRM. In fact, if more than one CRM is available, it is recommended to use more than one CRM to check the method over the range of values of the measurand relevant to the scope of the method.

9.3 Utilizing bias data

9.3.1 The bias estimates obtained from CRMs used in calibration can be used directly to make corrections. Such corrections can be additive, multiplicative, or a combination thereof. Knowledge concerning the calibration model and its particulars is necessary to decide how to apply a correction.

9.3.2 In testing, correction for bias is even more complicated, as the behaviour of the CRM(s) may not entirely reflect the behaviour of routine samples. In many cases, it is recommended to improve the method so that the bias is eliminated, rather than attempting to correct for it. Some standard test methods give criteria for acceptable bias.

NOTE If the behaviour of the CRM is inadequate and the measurement procedure cannot be improved, then the CRM is not suitable for use in assessing bias for the measurement procedure under investigation.

9.3.3 The expression for bias is given by

$$d = x_{\text{meas}} - x_{\text{CRM}} \tag{2}$$

and its associated standard uncertainty is evaluated using

$$u(d) = \sqrt{u_{\text{meas}}^2 + u_{\text{CRM}}^2} \tag{3}$$

9.3.4 If the bias is significant, that is, $|d| > U(d)$, usually attempts are made to find the cause of the bias and to reduce or eliminate it.

NOTE $U(d) = k \cdot u(d)$, where k denotes a suitably chosen coverage factor.

9.3.5 If sufficient reduction or complete elimination of the bias is not possible, the measurement result should be corrected for bias and the uncertainty associated with the bias should be included in the uncertainty evaluation. Corrections may be additive or multiplicative, depending on whether they depend on the value of the quantity to be corrected.

9.3.6 If the observed bias is not corrected for and significant, it should be included in the uncertainty budget. A rough approximation is to add the square of the bias (i.e. d^2) to the uncertainty budget to account for the uncorrected, significant bias.

9.3.7 If the bias is evaluated over a range of values for the measurand, an average bias can be calculated, with its associated uncertainty. A general approach is given in the GUM (ISO/IEC Guide 98-3:2008, F.2.4.5).

10 Calibration

10.1 General

10.1.1 For calibration, a CRM is needed. CRMs should be suitable for calibrating the equipment in terms of

a) physical form,

b) appropriateness of the property (properties) certified,

c) range of values and their relevance for the measurement range,

d) appropriateness of its reflection of the behaviour of routine samples (commutability).

10.1.2 It can be necessary to use a set of CRMs for instrument calibration, in particular if departure of proportional behaviour of the instrument reading with respect to the property value is possible.

10.1.3 The uncertainty associated with the property value should be used in evaluating the uncertainty of measurement due to calibration. For this purpose, the law of propagation of uncertainty of the GUM can be used, or any other mechanism for propagating probability distributions or uncertainties.

10.2 Establishing metrological traceability

10.2.1 The use of a CRM for calibration of an apparatus is a convenient way to establish metrological traceability for the calibration function obtained with this apparatus. Usually the property value(s) of the CRM is used in the calibration model.

10.2.2 In some cases a relevant CRM is only available as pure substance, whereas the calibration method requires another physical form. If this is the case, the value as obtained during calibrant preparation and its associated uncertainty should be used in the measurement process.

10.2.3 A certain degree of quality assurance should be applied when calibrating equipment with CRMs. At minimum, the calibration should be checked with a suitable QCM, a previously used calibrant, or by some other means showing that the last calibration is in agreement with previous calibrations.

NOTE Checking the consistency of calibrations may be combined with other quality assurance measures assuring the validity of measurement results.

10.3 Calibration models

10.3.1 The use of CRMs for calibration purposes is from the point of view of establishing metrological traceability and evaluating the uncertainty of measurement quite straightforward. The value assigned to the property of interest is entered in the calculation of the value assigned to the sample(s) measured. Annex B of this Guide describes for three commonly used cases how the value of the CRM enters in the calculations, namely

— single point calibration,

— bracketing,

— multipoint calibration.

10.3.2 Single point calibration is the simplest method; one calibrant (in this context the CRM) is used to calibrate the measurement equipment, which is then used to assign a value(s) to the sample(s) measured.

10.3.3 Bracketing requires two calibrants, one with a property value greater than the value(s) of the sample(s), and one with a property value smaller than those values. By means of linear interpolation between the two calibrants, values are assigned to other samples.

10.3.4 Multipoint calibration is widely used, in particular in analytical chemistry, to perform calibration of measurement equipment. A suite of calibrants is measured, and based on the responses measured, usually curvilinear regression is used to establish a relationship between the response measured and the quantity to be measured.

NOTE A simple form of a curvilinear relationship is a straight line.

10.3.5 Expressions for assigning a value on the basis of these three mainstream approaches and the associated uncertainty evaluation are given in Annex B.

11 Assigning values to other materials

11.1 General

11.1.1 Particularly in instrument calibration, CRMs are often used to prepare other RMs by means of mixing, dilution, or otherwise. The property value(s) for the newly prepared RM are partly based on the

property value(s) of the CRM used for preparation. These applications are covered under the generic heading “assigning values to other materials”. Methods of preparation include gravimetry and volumetry.

11.1.2 This application of a CRM is very frequently used. In fact, most of the calibrations performed in analytical chemistry are based on this role of a CRM. Pure materials are often used for preparing mixtures, or solutions which are in turn used for calibrating equipment. Sometimes, these mixtures or solutions are further diluted prior to use. The concentration, amount–of–substance fraction, or some other composition measure can be calculated on the basis of the purity data and the preparation data.

If the equipment used in the preparation process is calibrated appropriately, and the environmental conditions are monitored accordingly, it is possible to obtain property values that are metrologically traceable to SI.

NOTE 1 Environmental conditions can play a dominant role in the accuracy of gravimetry. In particular when weighing voluminous objects, such as gas cylinders, air buoyancy may play an important role. The need for controlling environmental conditions depends on the rigor of the uncertainty evaluation and the desired level of accuracy.

NOTE 2 Concentration measurements are, among others, temperature dependent. These effects may be small in view of other uncertainty components in a field laboratory, but they are not in view of other sources of uncertainty in the volumetric process.

11.1.3 Checking the consistency of values assigned to these calibrants is recommended. Such checks can be performed

— by comparing a new calibrant against an old, validated one,

— by assessing the effect of using the new calibrant as part of a quality control, by, e.g. measuring a QCM.

11.1.4 Comparison of a new calibrant against an old, validated one is necessary if the calibration has an important impact on the overall performance of the measurement procedure. The value calculated from the preparation of the new calibrant (x_{prep}) should be compared with that using the old calibrant (x_{meas}). The old, validated calibrant is used for calibrating the instrument used for the comparison. The new calibrant is validated if

$$\left|x_{\text{prep}} - x_{\text{meas}}\right| \leq k\sqrt{u_{\text{prep}}^2 + u_{\text{meas}}^2}$$

where k denotes a suitably chosen coverage factor at a 95 % level of coverage. In most cases, $k = 2$ is an appropriate choice (see ISO/IEC Guide 98-3:2008, Clause 7 for further guidance on choosing coverage factors).

Instead of applying the above criterion, the new calibrant can also be considered validated if the observed difference ($x_{\text{prep}} - x_{\text{meas}}$) is small in comparison with the uncertainty required by the measurement or test method.

11.2 Pure materials

11.2.1 Pure materials play a crucial role in establishing metrological traceability in many areas of measurement, in particular, but not limited to, chemistry.

NOTE “Pure” is an idealised concept, as is homogeneous or stable. In reality, no material is in an absolute sense pure.

11.2.2 For composition measurements, pure materials often form the basis of what is called the “traceability chain”, i.e. it is the first link between the pure material and composition measurements of the substance in question. Any material used for this purpose should have been characterized for impurities, and these should be identified and quantified as relevant for the intended use.

EXAMPLE 1 Nitrogen (6.0 grade) usually contains a few nmol/mol benzene. This impurity is not very relevant when making a synthetic natural gas mixture (where the amount–of–substance fraction of nitrogen is typically between 0,5 cmol/mol and 20 cmol/mol, and that of benzene, if at all, in the range of 1 µmol/mol to 10 µmol/mol), but it is highly relevant when making air quality standards (where the target amount fraction of benzene is 5 nmol/mol to 50 nmol/mol.

EXAMPLE 2 The presence of impurities in materials used for the fixed points on the temperature scale lead to deviations due to, e.g. freezing point depression.

11.2.3 Many chemicals and other pure materials come with data concerning the impurities. This information is only useful in a metrological context if it specifies

— units of measurement (e.g. mol/mol if expressed as amount–of–substance fractions),

— uncertainty associated with the assigned values.

EXAMPLE 1 For a testing laboratory analysing trace contaminants in soil, it usually suffices to compare a calibrant prepared with the new pure chemical against an old calibrant, from which it has been established that it is not significantly biased. Such establishment may be obtained from, e.g. the repeated analysis of a QCM, PTM (proficiency testing material), or using a CRM.

EXAMPLE 2 A manufacturer producing calibrants for testing laboratories should provide metrologically traceable composition data and hence perform an adequate identification and quantification of impurities of the pure materials (including the matrix).

11.2.4 Purity analysis is not limited to the pure materials being dissolved, diluted or otherwise made suitable for chemical composition measurement. The solvent, matrix gas, etc., should also undergo purity analysis, as it may also contain detectable quantities of impurities that may have an impact on the measurement results.

In measurement and testing, such purity analysis may be conducted as a reagent blank check, i.e. checking whether the level of impurities is sufficiently low to be neglected in subsequent steps in the measurement procedure.

11.3 Gravimetry and volumetry

11.3.1 Gravimetry and volumetry are widely used as techniques for preparing calibrants.

11.3.2 The property values of calibrants are calculated based on the procedure used to prepare them.

11.3.3 The uncertainty associated with the property values can be obtained using the law of propagation of uncertainty and the models indicated under 10.3.1.

11.3.4 Many calibrants are not stable with respect to one or more properties over time. In order to obtain valid results during the calibration of an apparatus, the property values assigned earlier to the calibrant should still be valid within their respective uncertainties. It can be necessary to perform some kind of stability checking.

11.3.5 Stability checking of calibrants can be performed in many ways. Some methods are

a) by conducting a stability test;

b) by comparing measurement results obtained using a new and old calibrant on the same sample, a retained PTM or a QCM;

c) by calibrating an instrument with the old calibrant and measuring the new one, followed by comparing the measured value against the value assigned to the calibrant;

d) check the entire measurement procedure with an independent CRM.

Stability testing can be laborious, but can be necessary if there are alternatives. Stability testing is covered in ISO Guides 34[1] and 35[2].

11.3.6 The values assigned to calibrants should be valid for their entire lifetime. Laboratories should set such lifetimes to their calibrants to make specific checking unnecessary and rely on good quality control to detect issues with calibrants.

If such lifetimes are to be determined, the quality of the calibrant should be checked regularly so that metrological traceability of the results obtained with these is assured, in particular if they are used over longer periods of time.

12 Conventional scales

12.1 General

12.1.1 Many measurement scales have been used since the earliest civilizations. Originally almost all of them were conventional, independent and inaccurate. Scientific and technical progress as well as international trade have led to both the need and the possibility of a unique, rational, self-consistent international system of units, the SI, which has been officially adopted worldwide. Nevertheless, it is not applicable to certain types of measurements for which it is necessary to create, sustain and use certain conventional units which are not within the scope of SI. In other cases the unit relating to the quantity to be measured lies within the frame of SI, but the reproduction of the unit according to the definition is technically difficult and expensive. It is therefore more convenient to carry out the measurement on a practical scale of reference values assigned to material properties. Though a reference value scale and a pure conventional scale differ theoretically from each other, they are similar with respect to the use of reference materials, and they will therefore be discussed together as conventional scales.

12.1.2 Conventional scales are based on the values assigned to reference materials. The assigned values are stated in standard specifications, international recommendations or other reference documents; therefore a reference material resulting in a fixed point on a conventional scale should have the same quality all over the world. CRMs of this type are certified for property values, i.e. they are measured on standard equipment with reference measurement.

12.1.3 It is evident that the CRMs ensure only the fixed points of a measurement scale. Measurement on a scale requires either a fixed point and a mathematical function passing through it, or two or more fixed points with stated means of interpolation between them.

NOTE Some special discontinuous scales exist, e.g. Mohs' scale for measuring hardness in geological tests. The scale is based on 10 minerals to which are assigned 10 grades of hardness; each harder mineral scratches the less hard one.

12.1.4 A conventional scale has two fundamental pillars: the certified reference material, providing the fixed point(s), and the standard specification (or similar document), giving the measurement procedure. Both of them should be strictly defined to ensure the compatibility of measurements on the conventional scale.

The standard specification provides detailed information necessary to establish and use a scale based on assigned values; or it may provide protocols for the experimental and calculation procedures to be used in measurements which depend on assumptions. It is advisable to specify the requirements of the certified reference material in the same standard specification as that in which the measurement procedure is described. By means of the necessary CRMs and relevant standard specifications, the user can create the measurement scale, and with the aid of such a scale can measure their sample or calibrate their instrument.

12.1.5 To evaluate the uncertainty of a measurement on the scale, the user should consider the uncertainties in the creation of the scale and the uncertainty associated with the determination of its fixed points by the CRM. Sometimes the users demand a level of uncertainty in the end-use which is lower

than the uncertainty of the fixed points defined by the CRM (e.g. in measurement of the pH of blood). They need to realize that the uncertainty of the measurements on the scale is necessarily greater than the uncertainty of the fixed points. The replicated measurement of the CRM and the setting up of a scale (the appropriate selection of the points, the characteristics and repeatability of the interpolating instrument, etc.) also contribute to the overall uncertainty.

12.1.6 The selection of CRMs for determining the fixed points of a scale should be directed by the required level of uncertainty of the end use. To minimize the uncertainty of the measured value on the scale, the user should employ CRMs which have been certified in terms of the units of the scale. Obviously, the user is expected to be familiar with all relevant information about the measurement procedure for creating the scale and the instructions for the correct use of the CRM.

In certain cases, the user can apply pure chemical compounds for determining the fixed points, if CRMs certified in the scale units are unavailable or expensive, or if their use is not necessary at the level of the uncertainty of the measurement. If this procedure is chosen, the user should be aware of the correlation between the purity of the material and the property on which the scale is based. The uncertainty of the measurement can be only roughly evaluated.

There is a great variety of conventional scales and the methods of application of the CRMs for determining them differ widely. Two examples are given in 12.2 and 12.3 to show some features of conventional scales.

12.2 pH-scale

12.2.1 Since absolute single-ion activities cannot be measured experimentally, it is recognized that the pH value is an inexact physical quantity. In order that measured pH be treated with as much significance as possible, a conventional pH scale has been adopted which is defined by reference solutions with assigned values of pH. These values have been determined by measuring the electromotive force (e.m.f.) of a hydrogen-silver/silver chloride cell without transference and by a given method of calculation, based on a convention.

12.2.2 Various national standard specifications describe the methods of preparing and assigning pH values to the reference solutions. The uncertainty of the certified values of these reference solutions is limited to a few thousandths of a pH unit.

12.3 Octane number

12.3.1 The octane-number scale is defined by ASTM-IP joint standard specifications. International Standards as well as a number of national standards refer to these documents. ASTM D 2699-95a/IP 237 and ASTM D 2700-95a/IP 236 describe the test methods for knock characteristics of motor fuels by the research method and by the motor method respectively. In both standards, the octane number of a fuel is determined by comparing its knocking tendency with those for blends of ASTM reference fuels of known octane number under standard operating conditions. The reference materials and blending accessories are given in annexes of both standards.

12.3.2 The ASTM standards refer to NIST SRM No. 1816a (iso-octane, purity 99,987 %) and SRM No. 1815a (*n*-heptane, purity 99,987 %). The principal use of these materials is in certifying the commercially produced ASTM Knock Test Reference Fuels. Specifications for these reference fuels are given in the standard, in which the suppliers are also listed. The responsibility for meeting the specifications for the reference materials rests with the suppliers. ASTM certification is based on the physical properties of the sample. Suppliers are required to test a sample of the reference material to be certified and at the same time test the corresponding SRM to provide traceability of production to an accepted reference material. A certificate is issued by ASTM to the suppliers, authorizing them to guarantee that the material shipped has been tested accordingly, and to quote the results of the tests.

13 Selection of CRMs and RMs

13.1 General

13.1.1 Reference materials, and certified reference materials in particular, can be used for various purposes in a measurement process. These purposes include:

— Calibration;

— Establishing metrological traceability;

— Method validation;

— Assignment of values to properties of other materials.

— Quality control.

13.1.2 It is highly recommended that RMs and CRMs produced (and for CRMs certified) in accordance with ISO Guides 34[1] and 35[2] are used. Ideally, such conformity is explicitly stated in the accompanying documentation. The user should check that this is the case, and if it is not stated, ask the producer of the RM.

13.1.3 Not every RM can be used for every purpose. Furthermore, a given RM can only be used for a single purpose in a specific measurement.

EXAMPLE A synthetic natural gas mixture certified for the contents of methane, ethane, propane, *iso*-butane, *n*-butane, nitrogen, and carbon dioxide can be used for calibration of a gas chromatograph (GC). It cannot, in the same measurement, simultaneously be used to check the trueness of the GC. The same CRM can however be used to assess other aspects of performance such as retention times, peak separation, and precision.

13.1.4 Reference materials can come in different forms[16]. Some of the common forms include:

a) **pure substances** characterized for chemical purity and other properties such as melting point, viscosity, enthalpy of combustion;

b) **standard solutions and gas mixtures**, often prepared gravimetrically from pure substances;

c) **matrix reference materials**, characterized for the composition of selected properties such as the content of specified chemical constituents. Such materials may be prepared from naturally occurring materials, or by synthesis;

d) **physical-chemical reference materials** characterized for properties such as melting point, viscosity, octane number, flash point hardness, and absorbance;

e) **reference objects or artefacts** characterized for functional properties such as taste, odour, etc. This type also includes specimens characterized for properties ranging from fibre type to microbiological specimens.

Whether a CRM or some other kind of RM is selected for a particular purpose depends on many factors, including the specific application, and the availability.

13.1.5 Reference materials can be found in various databases on the Internet, including COMAR, the key comparison database of the national metrology institutes (KCDB), and the websites of reference material producers.

13.2 Selection of a CRM

13.2.1 For the purpose of this Guide, it is understood that any CRM is accompanied by at minimum the following documentation:[17]

— general particulars of the producer or certifying body assuming responsibility for the certificate;

— a description of the material, including a clear statement of the species used for materials from animal or plant origin;

— intended use of the CRM;

— for each property, a value and its associated (expanded) uncertainty;

— the metrological traceability of the certified property values;

— expiry date (period of validity) of the certificate;

— instructions for use, including any limitations;

— appropriate storage conditions.

NOTE For CRMs intended for qualitative measurement, properties may be expressed as enumerated values (such as colours), and uncertainties may be expressed as probabilities.

13.2.2 The intended use of a CRM states the purposes for which the CRM can be used in a measurement process.

13.2.3 Laboratories should be able to explain and justify the basis of selection of all CRMs and of course any decision not to use a CRM. A formal suitability assessment should be performed by the user, unless it can be shown that the choice of RM will not significantly affect measurement results. The process of assessing the suitability of a CRM is visualized in Figure 2. The various aspects to be included in the assessment are given in 13.3.

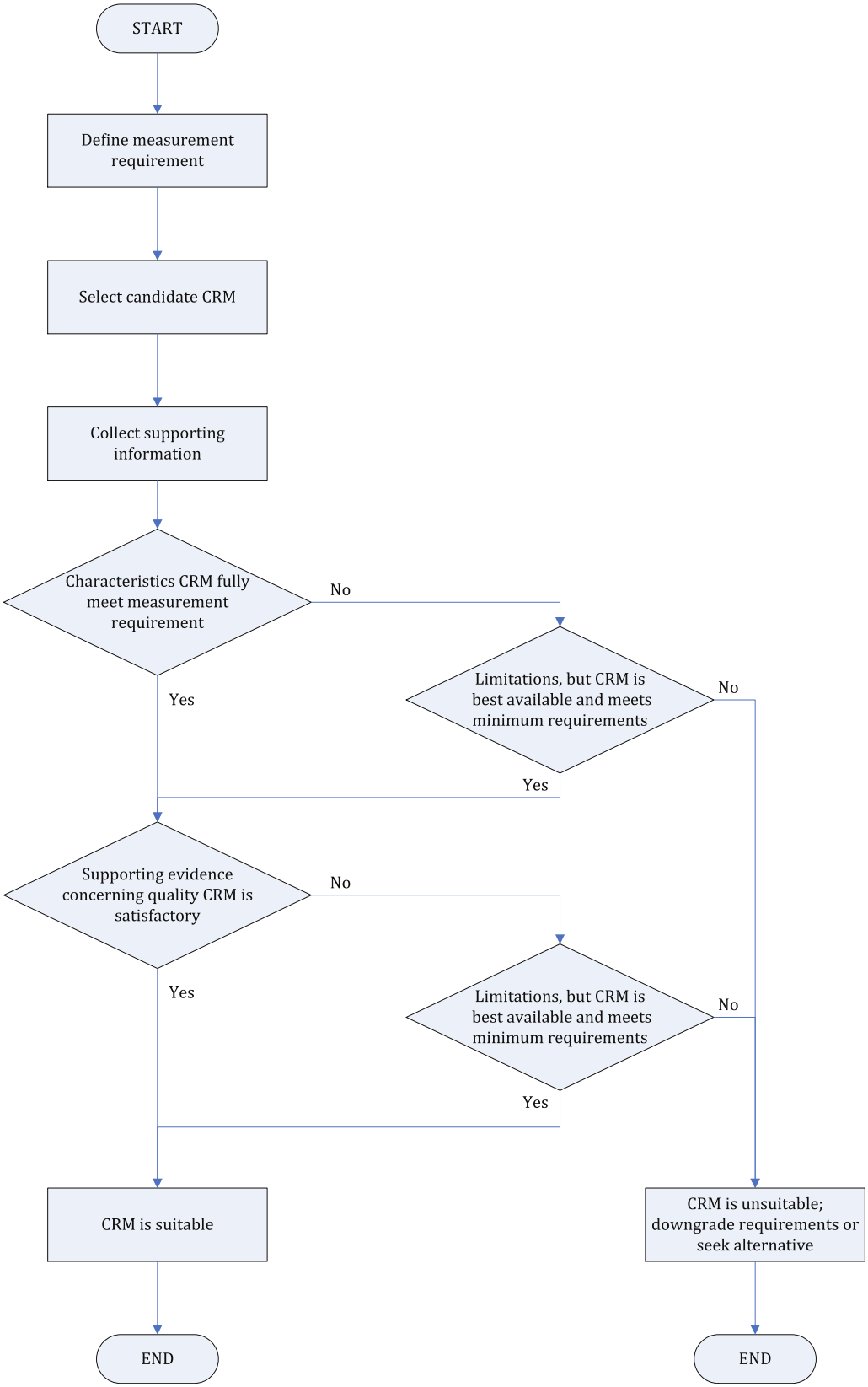

Figure 2 — Assessment of the suitability of a CRM (in modified form, from [16])

13.3 Selection of RMs

13.3.1 For the discussion of the selection of RMs (other than CRMs), there are two cases to be considered:

a) characterized RMs;

b) uncharacterised RMs, that is, RMs that come without a property value.

RMs that come with stated property values should meet the requirements of ISO Guide 35 and therefore these property values should be metrologically traceable (preferably to SI). In order to fulfil their purpose, these RMs should come with a documentation package that contains at least the same information as would otherwise be stated on an RM certificate.[17] For these RMs, the same considerations apply as for the selection of CRMs (see 13.2).

13.3.2 Many RMs come without stated property values. For most applications, the nominal values, or the range(s) within which the property value(s) are expected to lie, should be known. These RMs are often used for various kinds of precision control, such as day-to-day within laboratory quality control. Furthermore, in order to be useful, these RMs should have been checked for homogeneity and stability as described in ISO Guide 35.[2]

NOTE There are situations, where RMs without –for example– stability data can still be usefully applied in measurement processes. These situations are limited to cases, where the stability of the RM can be demonstrated indirectly by some other reference, such as including in the measurement process another RM (or CRM) from which the stability has been established.

13.4 Relevance to the measurement system

The user of the CRM should decide what properties of the CRM are relevant to the measurement procedure, taking into account the approach to certification, the statement on intended use and instructions for the correct use of the CRM on the certificate.

a) **Level.** The CRM should have properties at the level(s) appropriate to the level at which the measurement process is intended to be used, e.g. concentration.

b) **Matrix**. The CRM should have a matrix as close as possible to the matrix of the material to be subjected to the measurement process, e.g. carbon in low-alloy steel or carbon in stainless steel.

c) **Form**. It may be a test piece or a manufactured article or a powder. It may need preparation. It should be used in the same form (e.g. solid, gas, etc.) as the sample to be measured.

d) Minimum sample intake. Whenever the CRM certificate specifies a minimum amount that should be taken to obtain a sub-sample representative of the bulk, this specification should be adhered to.

e) **Quantity**. The quantity of the CRM should be sufficient for the entire experimental programme, including some reserve if it is considered necessary. Avoid having to obtain additional new units of the CRM later in a given measuring process, unless the CRM is provided as single use units.

f) **Stability.** Wherever possible the CRM should have stable properties throughout the experiment. Three situations can exist:

   1) the properties are stable and no precaution is necessary;

   2) the certified value of the properties may be influenced by storage conditions, in which case the container should be stored, both before and after its opening, in the way described on the certificate;

   3) the properties (which are changing at a known rate) at specific times are defined in a certificate supplied with the CRM. The user should obey the instructions for use as provided on the

certificate or associated documents. The property values and stated uncertainties are only valid under these conditions.

g) **Acceptable uncertainty of the certified value.** The uncertainty of the certified value should be compatible with the precision and trueness requirements outlined in 1.1.

h) **Commutability**. Where relevant, the user should assess whether the CRM is commutable with respect to the intended use. [18][19] Data from an assessment performed by the CRM producer may be available to aid assessing the commutability of the CRM by the user.

In this case, matrix effects and effects due to sample preparation should be evaluated

Annex A
(informative)

Key characteristics of a reference material with respect its common applications

See Table A.1.

Table A.1 — Key characteristics of reference materials and their relevance in common applications

	Precision control	Bias control	Calibration/ conventional scales	Assigning values to other materials
Specification of the property of interest	Required	Required	Required	Required
Property value		Required	Required	Required
Stated uncertainty		Required	Required	Required
Specified level of homogeneity	Required	a	a	a
Specified level of stability	Required	a	a	a
Statement of metrological traceability		Required	Required	Required
Instructions for use	Required	Required	Required	Required
Expiry date of the certificate		Required	Required	Required

a Uncertainty contribution included in the stated uncertainty associated with the property value.

Annex B
(informative)

Calibration models and associated uncertainty models

B.1 Single point calibration

B.1.1 Given response y_{CRM} and y_{sample} obtained from measuring the CRM and sample respectively, and the property value x_{CRM} of the CRM, the result of the sample can be calculated using

$$x_{\mathrm{sample}} = x_{\mathrm{CRM}} \frac{y_{\mathrm{sample}}}{y_{\mathrm{CRM}}} \tag{B.1}$$

The uncertainty associated with x_{sample} can be calculated from

$$u^2\left(x_{\mathrm{sample}}\right) = x_{\mathrm{sample}}^2 \left\{ \frac{u^2\left(x_{\mathrm{CRM}}\right)}{x_{\mathrm{CRM}}^2} + \frac{u^2\left(y_{\mathrm{CRM}}\right)}{y_{\mathrm{CRM}}^2} + \frac{u^2\left(y_{\mathrm{sample}}\right)}{y_{\mathrm{sample}}^2} \right\} \tag{B.2}$$

The uncertainty $u(x_{\mathrm{sample}})$ covers only effects from repeatability of measurement and the uncertainty associated with the property value of the CRM; this uncertainty budget should usually be complemented with uncertainty contributions from, e.g. sampling, sub-sampling and sample transformation.

B.1.2 The use of this model and associated uncertainty evaluation is limited to situations, where the relationship between the x and y variables can be modelled accurately over the range of interest using the relationship

$$y = a_1 x$$

where a_1 denotes the slope. If there is a significant zero deviation or there is nonlinearity to be considered, the single point calibration model may be invalid, or other terms should be included in the expression for the standard uncertainty.

B.2 Bracketing

B.2.1 Bracketing requires two CRMs, one in value above the value of the unknown sample, and one below this value. The results of the two CRMs are used for linear interpolation. Hence, the interval chosen should be small enough so that non–linearity of the detector (if any) does not lead to a bias in the value assigned to the unknown sample. Consequently, the maximum size of the interval depends on the non–linearity of the measurement system used.

B.2.2 The model for bracketing reads as follows

$$x_{\mathrm{sample}} = \frac{x_2 - x_1}{y_2 - y_1}\left(y_{\mathrm{sample}} - y_1\right) + x_1 \tag{B.3}$$

where y denotes the response and x the quantity to be measured (for example, a concentration). The values of the CRMs are denoted by x_1 and x_2, the respective responses by y_1 and y_2, the response of the sample by y_{sample}, and the value of the sample by x_{sample}. The combined standard uncertainty associated with x_{sample} can be expressed in terms of the uncertainties of the variables at the right-hand side by using the law of propagation of uncertainty.

The expression for the combined standard uncertainty associated with x_{sample} is calculated as follows

$$u^2\left(x_{\text{sample}}\right)=\left(\frac{\partial x_{\text{sample}}}{\partial x_1}\right)^2 u^2\left(x_1\right)+\left(\frac{\partial x_{\text{sample}}}{\partial x_2}\right)^2 u^2\left(x_2\right)+\left(\frac{\partial x_{\text{sample}}}{\partial y_1}\right)^2 u^2\left(y_1\right)+\left(\frac{\partial x_{\text{sample}}}{\partial y_2}\right)^2 u^2\left(y_2\right)$$
$$+\left(\frac{\partial x_{\text{sample}}}{\partial y_{\text{sample}}}\right)^2 u^2\left(y_{\text{sample}}\right)$$
(B.4)

where the sensitivity coefficients can be obtained from analytical differentiation of the expression for x_{sample}. The expressions for the sensitivity coefficients read as follows

$$\frac{\partial x_{\text{sample}}}{\partial x_1}=-\frac{y_{\text{sample}}-y_1}{y_2-y_1}+1$$

$$\frac{\partial x_{\text{sample}}}{\partial x_2}=\frac{y_{\text{sample}}-y_1}{y_2-y_1}$$

$$\frac{\partial x_{\text{sample}}}{\partial y_1}=\frac{x_2-x_1}{\left(y_2-y_1\right)^2}\left(y_{\text{sample}}-y_1\right)-\frac{x_2-x_1}{y_2-y_1}$$

$$\frac{\partial x_{\text{sample}}}{\partial y_2}=-\frac{x_2-x_1}{\left(y_2-y_1\right)^2}\left(y_{\text{sample}}-y_1\right)$$

$$\frac{\partial x_{\text{sample}}}{\partial y_{\text{sample}}}=\frac{x_2-x_1}{y_2-y_1}$$

The uncertainty $u(x_{\text{sample}})$ covers only effects from repeatability of measurement and the uncertainty associated with the property value of the CRMs; this uncertainty budget should usually be complemented with uncertainty contributions from, e.g. sampling, sub-sampling and sample transformation.

B.3 Multipoint calibration

B.3.1 Multipoint calibration is widely used as a calibration technique. It does not require assumptions such as no zero deviation (single point calibration), linearity of the relationship between instrument response y and property value x (bracketing and single point calibration).

B.3.2 Data from a multipoint calibration usually require some kind of fitting routine and an appropriate model in order to obtain a value for an unknown sample, given the measured response of this sample. In most cases, a least squares procedure should be used to obtain the parameters of the calibration model.

B.3.3 There exist a wide variety of approaches to curve fitting. In order to propagate the uncertainties associated with x and y, a method should be used that minimises squared differences in both directions, taking into consideration the respective uncertainties. In gas analysis, this method is documented as ISO standard[20].

B.3.4 As next best option for least squares fitting, the traditional ordinary least squares can be used. A worked example is given in Reference [23].

B.4 Standard addition

A specific kind of multipoint calibration is known in the analytical chemical literature as the standard addition method. Rather than using separate calibrants, the calibration standard is added to the (transformed) sample. By means of extrapolation, the value of the measurand is determined.

Annex C
(informative)

Decision errors

The assessment of a measurement process on the basis of precision and trueness is always subject to rendering an incorrect conclusion because of

a) the uncertainty of measurement results, and

b) the limited number of replicate results usually performed.

Increasing the number of measurements tends to decrease the chance of an incorrect conclusion but, in many instances, the risk of making a wrong conclusion has to be balanced in economic terms against the cost of increasing the number of measurements. Accordingly, the rigour of the criteria developed for assessing a measurement process must take into account the level of precision and trueness requisite for the end-use.

For the purposes of this Guide, the term "null hypothesis" is applied.

In this case, the null hypothesis is that the measurement process has bias no greater than the limit chosen by the experimenter and variance no greater than the predetermined value; the alternative hypothesis is the hypothesis which is opposed to the null hypothesis (see also ISO 3534-1[24]).

There are two types of possible error in accepting or rejecting the null hypothesis.

a) **Error type I:** The error committed in rejecting the null hypothesis when in reality the null hypothesis is true.

    1) **Type I risk:** The probability of committing error type 1. Its value varies according to the real situation.

    2) **Significance level:** The given value, usually designated by α, which limits the probability of committing error type 1.

b) **Error type II:** The error committed in failing to reject the null hypothesis when in reality the alternative hypothesis is true.

    1) **Type II risk:** The probability, usually designated by β, of committing error type II. Its value depends on the real situation and can be calculated only if the alternative hypothesis is adequately specified.

    2) **Power of test:** The probability of not committing error type II, usually designated by $(1 - \beta)$. It is the probability of rejecting the null hypothesis when in reality the alternative hypothesis is true.

The choice of the values of both α and β is usually based on economic considerations dictated by the importance of the consequences of the decision. These values, as well as the alternative hypothesis, should be chosen before the start of the measurement process.

Bibliography

[1] ISO Guide 34, *General requirements for the competence of reference material producers*

[2] ISO Guide 35, *Reference materials — General and statistical principles for certification*

[3] ISO Guide 80, *Guidance for the in-house preparation of quality control materials (QCMs)*

[4] ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories*

[5] ISO 15189, *Medical laboratories — Requirements for quality and competence*

[6] NEN 7779:2008, *Environment – Measurement Uncertainty*

[7] ISO 5725-1, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

[8] ISO 5725-2, *Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*

[9] ISO 5725-3, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

[10] ISO 5725-4, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method*

[11] ISO 5725-5, *Accuracy (trueness and precision) of measurement methods and results — Part 5: Alternative methods for the determination of the precision of a standard measurement method*

[12] ISO 5725-6, *Accuracy (trueness and precision) of measurement methods and results — Part 6: Use in practice of accuracy values*

[13] ISO 7870-5, *Control charts — Part 5: Specialized control charts*

[14] ISO 16269-4, *Statistical interpretation of data — Part 4: Detection and treatment of outliers*

15] Snedecor G.W., & Cochran W.G. *Statistical Methods (8th edition).* Ames, Iowa. Blackwell Publishing Professional

[16] International Laboratory Accreditation Cooperation. ILAC G9: 2005, *Guidelines for the Selection and Use of Reference Materials*. http://www.ilac.org

[17] ISO Guide 31, *Reference materials — Contents of certificates and labels*

[18] ISO 15194, *In vitro diagnostic medical devices — Measurement of quantities in samples of biological origin — Requirements for certified reference materials and the content of supporting documentation*

[19] Vesper H.W., Miller W.G., Myers G.L. Reference Materials and Commutability. Clin. Biochem. Rev. 2007, **28**(4), pp. 139-147

[20] ISO 6143, *Gas analysis — Comparison methods for determining and checking the composition of calibration gas mixtures*

[21] Analytical Methods Committee (AMC) Technical brief No. 10. *Fitting a linear functional relationship to data with error on both variables*. Royal Society of Chemistry 2002, UK

[22] Ripley B.D., & Thompson M. Regression techniques for analytical bias. Analyst (Lond.). 1987, **112** pp. 377–383

[23] Ellison S.L.R., Rosslein M., Williams A. eds. Eurachem/Citac Guide — Quantifying uncertainty in analytical measurement. LGC, London, Third Edition, 2012

[24] ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ICS 71.040.30

Price based on 31 pages