GUIDE 33

Uses of certified reference materials

Second edition 2000

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

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.ch
Web www.iso.ch

Printed in Switzerland

Contents

Page

Foreword iv

Introduction iv

1 Scope 1

2 Terms and definitions 1

3 Symbols and subscripts 4
3.1 Symbols 4
3.2 Subscripts 5

4 Statistical considerations 5
4.1 Basic assumptions 5
4.2 Decision errors 5

5 The role of certified reference materials in measurement science 6
5.1 General 6
5.2 The role of certified reference materials in the storage and transfer of information on property values 6
5.3 Use of certified reference materials for measurement traceability 7
5.4 The role of reference materials in the International System of units (SI) 7
5.5 Defining and creating conventional scales 8

6 Assessment of a measurement process 10
6.1 The cases to be considered 10
6.2 Requirements of limits 10
6.3 Choice of CRM 11
6.4 Experimental procedure 12
6.5 General remarks 22

Bibliography 23

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.

Guides are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

Draft Guides adopted by the responsible Committee or Group are circulated to national bodies for voting. Publication as a Guide requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this Guide may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO Guide 33 was drawn up by the ISO Committee on reference materials (REMCO) and was approved by ISO member bodies.

This second edition cancels and replaces the first edition (ISO Guide 33:1989), which has been technically revised.

Introduction

Today's world of modern technology requires a large number of certified reference materials (CRMs) in widely diverse fields, and the demand for such materials is expected to increase. The preparation of a CRM is a time-consuming, meticulous and expensive endeavour and consequently it has not always been, and will continue not to be, possible to satisfy the demand for all types and quantities of CRMs. For this reason, CRMs should be used properly, i.e. effectively, efficiently and economically.

Certified reference materials should be used on a regular basis to ensure reliable measurements. However, in doing so, the magnitude of the supply of that CRM, its relative cost, its availability (accessibility) and the measurement technique, be it destructive or non-destructive, should be considered. Also important to the user is the fact that the misuse of a CRM may not provide the intended information.

Misuse of CRMs differs from incorrect use. The user of a CRM is expected to be familiar with all information pertinent to the use of the CRM as specified in its certificate. He should comply with such factors as the period of validity of the CRM, the prescribed conditions for storage of the CRM, instructions for the use of the CRM, and specifications for validity of the certified properties of the CRM. A CRM should not be used for a purpose other than that for which it was intended. Nevertheless, from time to time, when a user must resort to applying a CRM in an incorrect manner because of the unavailability of a suitable CRM, he should be fully cognizant of the potential pitfalls and therefore assess his measurement output accordingly.

There are many measurement processes in which CRMs are in general use but are replaceable by a host of working standards, such as homogeneous materials, previously analysed materials, pure compounds, solutions of pure elements, etc. Some examples are where only a "rough" estimate of the trueness or precision of a method is sought, where "blind" unknown check samples are used routinely in quality control programmes, and where only the variation in trueness or precision of a method with some parameter such as time, analyst, instrument, etc., is being evaluated. The first example illustrates the use of a CRM in which the well-defined certified value and uncertainty of the CRM is under-utilized. The others illustrate the case in which a series of "one-time" trueness and precision assessments are compared with one another. There is no need to base that comparison on a well-defined certified value and uncertainty of a CRM. The advantages in using CRMs are that the user has the means to assess the trueness and precision of his measurement method and establishes metrological traceability for his results.

Whether the use of CRMs in these procedures is in fact "misuse" depends largely on the availability and relative cost of the CRMs. Where CRMs are in short supply or very expensive, their use would indeed be misuse. However, for CRMs in ample supply or where similar CRMs are available from one or more sources, it is strongly recommended that CRMs always be used instead of in-house standards because of the resultant enhanced confidence in the measurement output.

It is important that users remain aware that the preparation of in-house standards for use instead of CRMs has an associated cost based on factors such as material cost, facility usage charges, personnel labour rates, etc., in which the material cost is in general the lowest. For some CRMs such as the complex compositional materials certified for chemical composition, the cost of preparing in-house standards to match the composition of real samples can exceed that of available CRMs. In these cases, the use of CRMs is recommended.

The user should be aware of the potential misuse of CRMs as "blind" unknown check samples in quality control programmes. Where there are only a few CRMs in an area of expertise, they are easily recognized and they may therefore not satisfy the intended purpose. Moreover, the same CRM should never be used for both calibration purposes and as "blind" unknown check sample in a measurement process.

The misuse of CRMs can also occur when the user does not fully take into account the uncertainty in the certified property. The combined standard uncertainty of a certified property of a CRM can have contributions from the inhomogeneity of the material, the within-laboratory uncertainty and, where applicable, the between-laboratory uncertainty. The level of homogeneity defined for a CRM by the producer is dependent on both the statistical design used to evaluate it and the repeatability of the method of measurement. For certain CRMs, the level of homogeneity is valid for a test portion defined by mass, physical dimension, time of measurement, etc. The user should be aware that the use of a test portion that does not meet or exceed that specification could severely increase the contribution of the inhomogeneity of the CRM to the uncertainty of the certified property, to the point where the statistical parameters of certification are no longer valid.

The variation in the repeatability of different methods has another implication for the user. Since the degree of inhomogeneity of a CRM is dependent on the repeatability of the method of measurement, it is possible that a user, in applying a method capable of better repeatability, could detect inhomogeneity in that CRM. In such cases, the observed inhomogeneity is already accounted for in the statistical parameters for the certified property and therefore the statistical tests presented in this Guide remain valid, but the scientific basis for using that particular CRM to give a true assessment of the user's method should again be questioned.

It is well known that different methods of measurement of a property are not capable of equal repeatability. Accordingly there could arise instances where the user may wish to assess a method that has greater repeatability than that or those used in the certification of the CRM. In such cases, the statistical tests presented in this Guide remain valid but the scientific basis for using that particular CRM to give a true assessment of the precision (and possibly the trueness) normally expected from the user's method should be questioned. It is recommended that the user resorts to a CRM of lesser uncertainty, if available.

For CRMs certified by a primary method, the user should not assume that his method is capable of matching the precision and trueness reported for the CRM. It is unreasonable therefore to apply the statistical procedures in this Guide for assessing the trueness and precision of a method by application to a CRM using the certification parameters for a property reported in the certificate. The user, as a consequence, should either experimentally establish or make estimates based on available information for those parameters that are more appropriate. Similarly, where a user applies a method to a CRM that has been certified by a single different method, the user should not assume that the certification parameters for the certified property are applicable to his method except in cases where the trueness and precision capable by both methods are known to be comparable.

One of the important considerations in selecting a CRM for use, either in assessing the trueness and precision of a method or in the calibration of instruments in a method, is the level of uncertainty required by the end-use of the method. Obviously the user should not apply a CRM of greater uncertainty than permitted by the end-use.

The selection of CRMs should take into account not only the level of uncertainty required for the intended purpose but also their availability, cost, and chemical and physical suitability for the intended purpose. For example, the unavailability or high cost of one CRM could force a user to resort to using another CRM of greater uncertainty than the preferred one. Also, in chemical analysis, a CRM of greater, but still acceptable, uncertainty in the certified property may be preferred over another CRM because of better matching with the composition of real samples. This could result in minimizing "matrix" or chemical effects in the measurement process which are capable of causing errors far greater than the difference between the uncertainties of the CRMs.

In conclusion, CRMs are meant to fulfil many purposes. Accordingly, a CRM used properly for one purpose in one laboratory may be misused for another purpose in another laboratory. It is recommended that the user consider the suitability of a CRM for his intended purpose on a case-by-case basis.

Uses of certified reference materials

1 Scope

This Guide discusses the uses of certified reference materials (CRMs) and their correct applications.

Clause 2 of this Guide presents definitions (with indication of their sources) of terms used, and clause 4 sets out the statistical considerations on which the Guide is based.

Clause 5 discusses the role of CRMs in measurement science and in realization of conventional measurement scales.

Clause 6 presents recommendations for developing criteria for the assessment of the precision and trueness of a measurement procedure by the use of CRMs. It pertains only to CRMs characterized to be homogeneous as described in ISO Guide 35 [4].

NOTE The use of CRMs is essential for assessment of trueness and optional for assessment of precision.

This Guide does not describe the use of certified reference materials as calibrants. This subject is treated in ISO Guide 32 [3].

2 Terms and definitions

For the purposes of this Guide, the following terms and definitions apply.

2.1
measurement process
all the information, equipment and operations relevant to a given measurement

NOTE This concept embraces all aspects relating to the performance and quality of the measurement; it includes, for example, the principle, method, procedure, values of the influence quantities and the measurement standards.

[VIM:1993]

2.2
influence quantity
quantity that is not the measurand but that affects the result of the measurement

EXAMPLE Ambient temperature; frequency of an alternating measured voltage.

[VIM:1993]

2.3
reference material
RM
a material or substance one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to materials

[ISO Guide 30:1992]

2.4
certified reference material
CRM
a reference material, accompanied by a certificate, one or more of whose property values are certified by a procedure which establishes traceability to an accurate realization of the unit in which the property values are expressed, and for which each certified value is accompanied by an uncertainty at a stated level of confidence

[ISO Guide 30:1992]

2.5
precision
the closeness of agreement between independent test results obtained under prescribed conditions

[ISO 3534-1]

2.6
repeatability
precision under repeatability conditions

[ISO 3534-1]

2.7
repeatability conditions
conditions where independent test results are obtained with the same method on identical test material in the same laboratory by the same operator using the same equipment within short intervals of time

[ISO 3534-1]

2.8
repeatability standard deviation
the standard deviation of test results obtained under repeatability conditions

NOTE It is a measure of the dispersion of the distribution of test results under repeatability conditions.

[ISO 3534-1]

2.9
repeatability limit
r
the value less than or equal to which the absolute difference between two single test results obtained under repeatability conditions is expected to be with a probability of 95 %

[ISO 3534-1]

2.10
reproducibility
precision under reproducibility conditions

[ISO 3534-1]

2.11
reproducibility conditions
conditions where test results are obtained with the same method on identical material in different laboratories by different operators using different equipment

[ISO 3534-1]

2.12
reproducibility limit
R
the value less than or equal to which the absolute difference between two single test results obtained under reproducibility conditions is expected to be with a probability of 95 %

[ISO 3534- 1]

2.13
bias
the difference between the expectation of the test results and an accepted reference value

NOTE Bias is a systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.

[ISO 3534-1]

2.14
accuracy
the closeness of agreement between a test result and the accepted reference value

NOTE The term accuracy, when applied to a set of test results, involves a combination of random components and a common systematic error or bias component.

[ISO 3534-1]

2.15
trueness
the closeness of agreement between the average value obtained from a large series of test results and an accepted reference value

NOTE The measure of trueness is usually expressed in terms of bias.

[ISO 3534-1]

2.16
uncertainty
<of measurement> parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand

[VIM:1993, GUM:1993]

NOTE This definition is that of the *Guide to the Expression of Uncertainty in Measurement* (GUM) in which its rationale is detailed (see, in particular, 2.2.4 and annex D [5].

2.17
estimation
the operation of assigning, from the test result in a sample, numerical values to the parameters of a distribution chosen as the statistical model of the population from which this sample is taken

[ISO 3534-1]

2.18
estimate
the result of an estimation

[ISO 3534-1]

2.19
estimator
a statistic used to estimate a population parameter

[ISO 3534-1]

2.20
null hypothesis
the hypothesis to be rejected or not rejected (accepted) at the outcome of the test

[ISO 3534-1]

3 Symbols and subscripts

3.1 Symbols

a_1; a_2	adjustment values chosen in advance
$E(x)$	expectation of a random variable
G	Grubbs' test statistic
n	number of replicate results
p	number of laboratories participating in interlaboratory measurement programme
r	repeatability limit
R	reproducibility limit
s	estimate of a standard deviation
$V(x)$	variance of a random variable
x	measurement result
$\overline{x}$	arithmetic mean of measurement results
$\overline{\overline{x}}$	overall (grand) mean of measurement results
α	significance level
β	type II error probability
δ	estimated bias of the measurement process
μ	accepted reference value of a property
ν	number of degrees of freedom
σ	true value of a standard deviation
σ_D	uncertainty of the measurement process expressed by its standard deviation

$\chi^2_{(n-1);\,0{,}95}$ 0,95th quantile of the χ^2 – distribution with $(n-1)$ degrees of freedom

3.2 Subscripts

c computed value

i identifier for an individual result

L between-laboratory (at CRM certification)

Lm between-laboratory (at the assessed method)

w within-laboratory

wo within-laboratory, required

4 Statistical considerations

4.1 Basic assumptions

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. They may be different if there is deviation from normality.

4.2 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[7]).

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.

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

— **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.

— **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.

— **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.

5 The role of certified reference materials in measurement science

5.1 General

Metrology is the field of knowledge concerned with measurement. Metrology or measurement science includes all aspects both theoretical and practical with reference to measurements, whatever their level of accuracy, and in whatever fields of science or technology they occur. This clause describes the role of reference materials in quantitative measurements.

5.2 The role of certified reference materials in the storage and transfer of information on property values

By definition (2.3), a reference material has one or more properties, the values of which are well established by measurement. Once the property value(s) of a particular CRM have been established, they are "stored" by the CRM (up to its expiration date) and are transferred when the CRM itself is conveyed from one place to another. To the extent that the property value of a CRM can be determined with a well-defined uncertainty, that property value can be used as a reference value for intercomparison or transfer purposes. Hence CRMs aid in measurement transfer, in time and space, similar to measuring instruments and material measures.

A CRM must be suitable for the exacting role it performs in storing and transferring information on measured property values. The following technical criteria (legal or commercial criteria may be relevant also) apply to the fitness-for-purpose of CRMs in general:

a) the CRM itself and the property value(s) embodied in it should be stable for an acceptable time-span, under realistic conditions of storage, transport and use;

b) the CRM should be sufficiently homogeneous that the property value(s) measured on one portion of the batch should apply to any other portion of the batch within acceptable limits of uncertainty; in cases of inhomogeneity of a large batch, it may be necessary to certify each unit from the batch separately;

c) the property value(s) of the CRM should have been established with an uncertainty sufficient to the end use(s) of CRM;

d) clear documentation concerning the CRM and its established property value(s) should be available. The property value(s) should have been certified, so the documentation should include a certificate, prepared in accordance with ISO Guide 31.

Whenever possible, the measurement of a given property value should have been made by an acceptable method having negligible uncertainty relative to end-use requirements and by means of measuring instruments or material measures which are traceable to national measurement standards. Subsequent use of a CRM with traceable property values ensures that traceability is propagated to the user. Since most national measurement standards