Eurachem

CITAC

Cooperation on International Traceability in Analytical Chemistry

! "#$%&'( )*)+,-.+)/"01'))

!"#$%&$ '(#)*+,'*-$.'&%)/+*,%'$,'$ 0%/12,+'(#$ 3""#""/#'*

2'%341)!10503)6768)

Use of uncertainty information in compliance assessment

English edition

Second edition 2021

ISBN 978-0-948926-38-9

Contents

Foreword

At the time of the first edition, work on compliance assessment had been carried out in other areas, particularly engineering, for the testing of electrical and mechanical products and the document followed the principles set out in ASME B89.7.3.1-2001.

"Compliance" and "conformity" are closely related terms. ISO often uses the term "conformity assessment"; ASME considers "conformance to specifications". Conformity assessment can, however, include a broad range of activities, from product testing to inspection and licensing. This Eurachem Guide on compliance assessment is primarily concerned with whether a measurement result complies with permitted limits, e.g. specifications, tolerances, regulatory or legal limits. The Guide accordingly uses the terms " compliance" or "compliance assessment" in relation to decisions about compliance with stated limits. In ISO/IEC 17025, compliance of a measurement result with stated limits is often used as the basis for a "statement of conformity".

This edition has been amended to take into account the developments in, e.g. *Guidelines on Decision Rules and Statements of Conformity* (ILAC G8) and *Evaluation of measurement data – The role of measurement uncertainty in conformity assessment* (JCGM 106).

Major changes in the second edition are:

- a list of abbreviations and symbols is added;
- the idea of an acceptance limit is introduced;
- decision rules that provide for conditional or inconclusive results are introduced (sometimes called "non-binary" decision rules);
- use of the lognormal distribution is introduced for some asymmetric cases;
- an Annex C is added introducing global and specific risks.

Abbreviations and symbols

The following abbreviations, acronyms and symbols occur in this Guide.

ASME	American Society of Mechanical Engineers
BIPM	International Bureau of Weights and Measures
CITAC	Cooperation on International Traceability in Analytical Chemistry
GUM	Guide to the Expression of Uncertainty in Measurement
JCGM	Joint Committee for Guides in Metrology
IEC	International Electrotechnical Commission
ISBN	International Standard Book Number
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
VIM	International Vocabulary of Metrology

g	Guard band
k	Coverage factor
n	Number of measurements
exp	Exponential function; $\exp(x) = e^x$
P	Probability (%) for compliance or non-compliance
L	Permitted limit for compliance
L_l	Permitted lower limit
L_u	Permitted upper limit
s	Standard deviation
s_G	Standard deviation of $\log_e$ data
x_i	Measured value
u	Standard uncertainty
u_{rel}	Relative standard uncertainty
^FU	Uncertainty factor
U	Expanded uncertainty

1 Introduction

In order to utilise a result to decide whether it indicates compliance or non-compliance with a specification, it is necessary to take into account the measurement uncertainty. Figure 1 shows typical scenarios arising when measurement results, for example the concentration of an analyte, are used to assess compliance with an upper specification limit. The vertical lines show the expanded uncertainty interval $\pm U$ on each measured value and the associated curve indicates the inferred probability density function for the value of the measurand, showing that there is a larger probability of the value of the measurand lying near the centre of the uncertainty interval than near the ends. Cases (i) and (v) are reasonably clear; the measurement results including the uncertainties provide good evidence that the value of the measurand is well above or well below the limit, respectively. In case (ii), however, there is a high probability that the value of the measurand is above the limit, but the limit is nonetheless within the uncertainty interval. Depending on the circumstances, and particularly on the risks associated with making a wrong decision, the probability of an incorrect decision may be or may not be sufficiently small to justify a decision of compliance. Similarly, in case (iv) the probability that the value of the measurand is below the limit may or may not be sufficient to take the result to justify a decision of compliance. In case (iii), the probability of an incorrect decision is 50 %. Without further information, which has to be based on the risks associated with making a wrong decision, it is not possible to use these three results, cases (ii), (iii) and (iv) to make a decision on compliance.

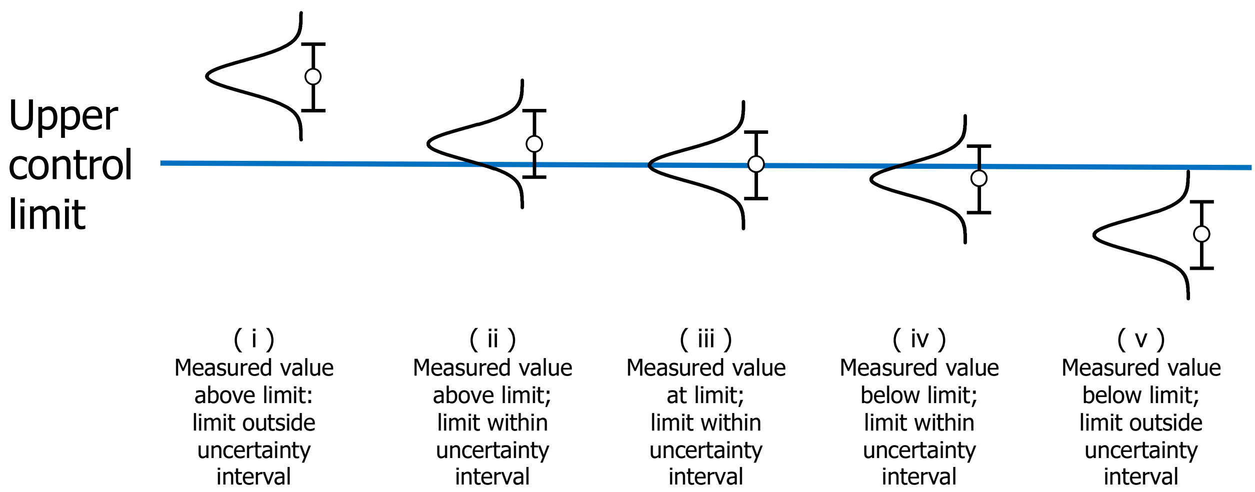

Figure 1 ─ Assessment of compliance with an upper limit

NOTE: This page is intentionally empty.

2 Scope

This document provides guidance on setting appropriate criteria for unambiguous decisions on compliance given results with associated uncertainty information. The key to the assessment of compliance is the concept of "decision rules". These rules give a prescription for the acceptance or rejection of an item based on the measured value, its uncertainty and the specification limit or limits, taking into account the acceptable level of the probability of making a wrong decision.

This document does not consider cases involving decisions based on multiple measurands. Some applications on compliance of multiple measurands can be found in references [1, 2].

When the decision on compliance is applied to the tested lot or batch of a substance or material, the measurement uncertainty component arising from the sampling could be important. This guide assumes that where the measurand implies a sampling requirement, the uncertainty includes components arising from sampling [3].

3 Definitions

Terms used in this guide generally follow the *International vocabulary of basic and general terms in metrology* (VIM) [4], the *Guide to the expression of uncertainty in measurement* (GUM) [5] and ILAC G8 [6]. Additional terms are taken from ASME B89.7.3.1-2001.1 [7]. A summary of the most important definitions used in this document is provided in Annex D – Definitions.

NOTE: This page is intentionally empty.

4 Decision rules

4.1 General

The key to the assessment of compliance is the concept of "Decision rules". These rules give a prescription for the compliance or non-compliance with a specification limit, taking into account the acceptable level of the probability of making a wrong decision. ISO/IEC 17025 defines a decision rule as: *rule that describes how measurement uncertainty is accounted for when stating conformity with a specified requirement* [8]. ISO/IEC 17025 also requires that, where relevant, the decision rule to be used should be agreed with the customer. ILAG G8 [6], JCGM 106 [9], Eurolab Report 1/2017 [10] and WADA TD2019DL [11] give an overview of decision rules and conformity with requirements. On the basis of a decision rule, an "acceptance zone" and a "rejection zone" can be determined, such that if the measurement result lies in the acceptance zone the item is declared compliant and if in the rejection zone it is declared noncompliant. The limits of the acceptance zone are called "acceptance limits".

A decision rule should have a well-documented method of determining the location of acceptance and rejection zones, ideally including acceptable levels of probability, P, that the value of the measurand 1) lies within the specification limit or 2) lies outside the specification limit.

Here, case 1) corresponds to high confidence of correct acceptance and a low probability of false acceptance, while case 2) provides high confidence of correct rejection and a low probability of false rejection.

The determination of the acceptance/rejection zone will normally be carried out by the laboratory when applying the decision rule.

The decision rule may also give:

- the maximum allowed uncertainty at the limit;
- an assumed distribution, e.g. normal or lognormal (see further Annex A);
- rules for rounding or truncation of measured values before assessing compliance;
- the required number of replicate measurements (if any) and the procedure for using replicate results including (for example) whether results should comply individually, or should be averaged before comparison with limits;
- procedures for dealing with outliers;
- procedures for further action, for a non-binary decision rule, when the decision is conditional (pass/fail);
- the procedure to be followed in the event of a conditional result requiring additional measurements;
- recommendations on how to report compliance/non-compliance, e.g. pass/fail, within tolerance/out of tolerance, within specification/out of specification;
- recommendations on how to state the decision rule used in the statement of compliance.

4.2 Decision rule with pass/fail using simple acceptance

For many situations, the decision rule is arranged to give a conclusive compliance decision: a pass or fail. The simplest example uses the specification limit as the acceptance limit, so that a result inside the limit is considered compliant. This is called "simple acceptance" or "shared risk" [6]. This corresponds to acceptance and rejection zones as shown in Figure 2. If the measurement result lies in the acceptance zone the item is declared compliant (pass) and if in the rejection zone it is declared non-compliant (fail). Referring to Figure 1 with an upper limit, cases (iv) and (v) are in the acceptance zone and cases (i) and (ii) are in the rejection zone. Case (iii) is in most cases regarded to be in the acceptance zone. To use this rule, there is usually a requirement that the measurement uncertainty has been considered and judged to be acceptable in order to keep the risk of making a wrong decision acceptable. However, for measured values close to a limit there is a risk (up to 50 %) of an incorrect decision using simple acceptance – see further Annex C – Producer and consumer risk.

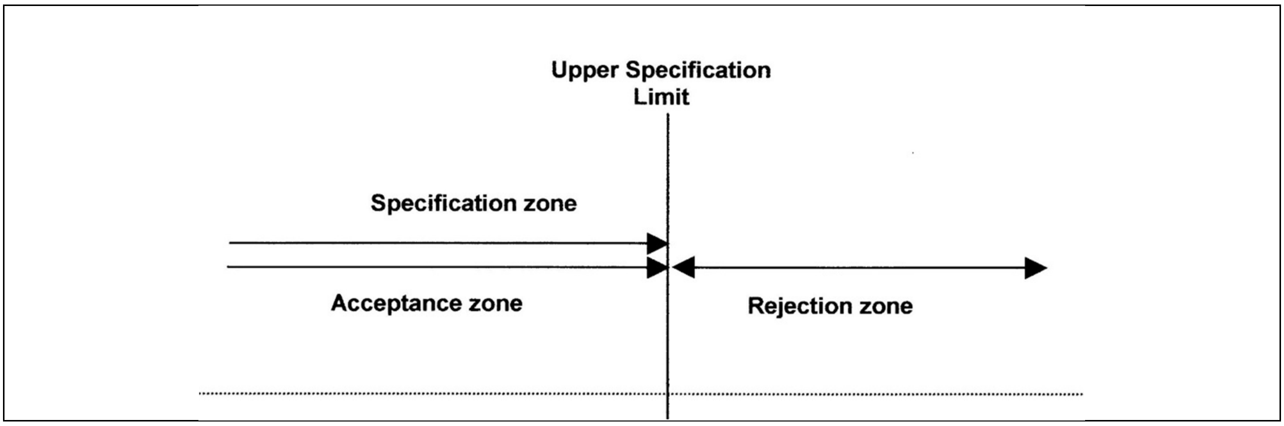

Figure 2 ─ Acceptance and rejection zones for simple acceptance with an upper limit. The acceptance limit is equal to the specification limit

4.3 Decision rule with pass/fail using guard band

For measured values very close to the limit, or when the uncertainty is large, simple acceptance can lead to an unacceptably high risk of an incorrect decision. Often, there is a need to be more confident in accepting or rejecting a test item. For these situations, acceptance and rejection zones can be determined as shown in Figure 3. For example, in Figure 3b, a measured value inside the acceptance zone is very unlikely to arise for a non-compliant test item. The interval g between the limit and the end of the acceptance zone is called a "guard band", reducing the risk of an incorrect decision. The use of guard bands provides a particularly simple way of defining decision rules; choosing the size of the guard defines an acceptance zone which can be used for decision making. In general, the guard band g will be a multiple of the standard uncertainty u. In the case when the relative uncertainty is less than about 15 to 20 % and the measured value is exactly at an upper or a lower acceptance limit, a guard band equal to $1.64u$ will give a probability of incorrect decision, α, of 5 % and a guard band equal to $2.33u$ implies a probability α of 1 %. The multiple can also be set explicitly; for example the guard band can be set to $2u$, i.e. the expanded uncertainty, U, is used as a guard band giving a probability of incorrect decision, α, of 2.5 %. A guard band can also be set to zero, $g = 0$. This is called simple acceptance or "shared risk", see Section 4.2. Determining the size of the guard band is discussed in Annex A and the risks of false rejection and acceptance respectively are dealt with in Annex C.

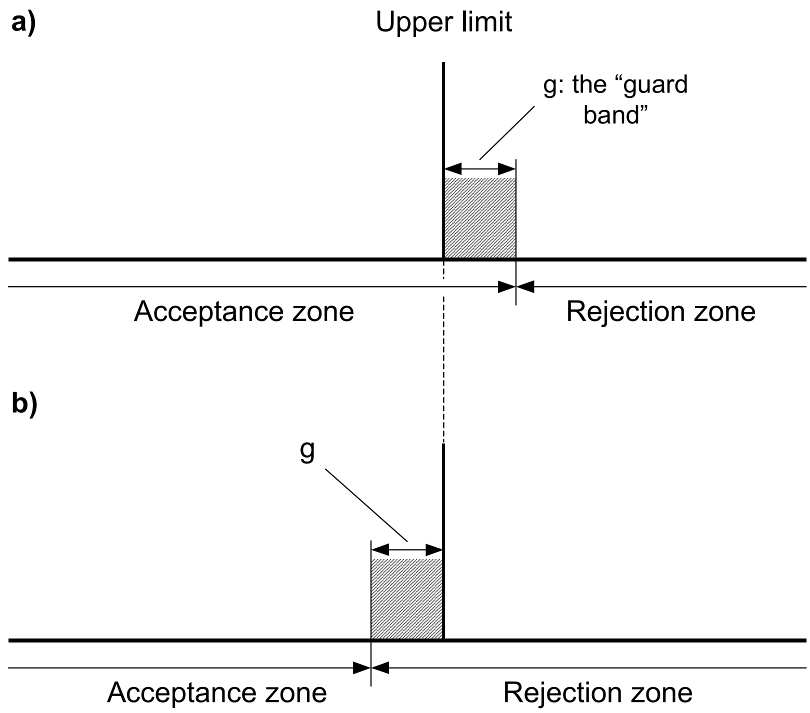

Figure 3 ─ Acceptance and rejection zones for an upper limit. The figure shows the relative positions of the acceptance and rejection zones for a) high confidence of correct rejection; b) high confidence of correct acceptance. The interval *g* is called the *guard band*. The upper end of the acceptance zone is the acceptance limit

In some cases the specification sets upper and lower limits, for example to control the composition of a product. Figure 4 shows the acceptance and rejection zones for such a case, where the guard bands have been chosen so that for a sample that is in compliance there is a high probability that the measurand is within the specification limits; that is high confidence of correct acceptance.

4.4 Decision rules with conditional or inconclusive results

Some decision procedures can include the possibility of a "conditional" pass/fail or an "inconclusive" outcome, typically where the specification limit is within the expanded uncertainty interval (see cases (ii), (iii) and (iv) in Figure 1). ILAC G8 [6] refers to this as a "non-binary" decision rule. For example, in Figure 1, case (ii) might be regarded as a "conditional fail" and cases (iii) and (iv) might be regarded as "conditional pass". Alternatively, a decision rule might choose to label these three cases as "inconclusive".

In many cases, a decision rule will provide for further testing in the event of a conditional or inconclusive result.

ILAC G8 includes additional examples of conditional acceptance and rejection.

4.5 Decision rule specifying a two stage procedure

To reduce the risks of false acceptance and/or rejection, some decision rules adopt a two-stage procedure, in which further measurements are made in the event of an inconclusive result. A general two stage procedure of this kind is described in ISO [12].

Such procedures need not use the same measurement procedure at each stage. For example, in order to reduce the cost of compliance assessment, a lower cost measurement can be performed first with a comparatively large uncertainty (this is often referred to as a “screening” test). If the initial result is inconclusive or close to a limit, a confirmatory procedure is applied to produce results with a smaller uncertainty. Most test items are then decided with adequate confidence at low cost, while borderline cases are decided with a more costly test with higher confidence.

The probabilities of false acceptance and/or false rejection in multi-stage procedures depend on the particular steps chosen, and are more complex than for single stage procedures. Probabilities for multi-stage procedures are outside the scope of the present Guide.

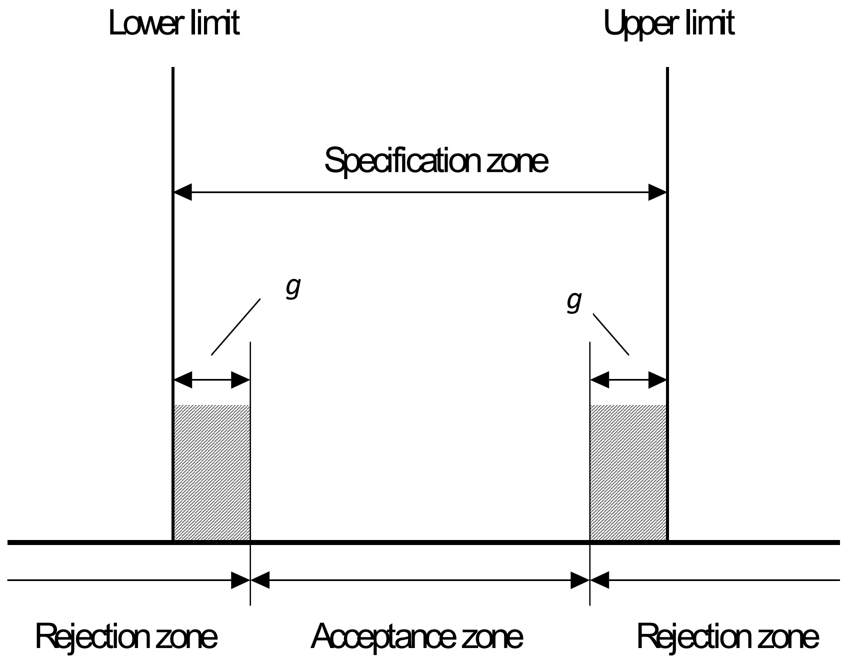

Figure 4 ─ Acceptance and rejection zones for a specification zone. The figure shows the relative positions of the specification limits and of the acceptance and rejection zones for high confidence of correct acceptance

5 Choosing acceptance and rejection zone limits

The size of the guard band *g* is chosen to meet the requirements of the decision rule and based on the value of the standard uncertainty *u* obtained from measurement close to the limit *L*. For example if the decision rule states that for non-compliance, the measured value should be greater than the limit plus 2*u*, then the size of the guard band is 2*u* giving an acceptance limit of $L + 2u$.

Similarly, if the decision rule is that there should be at least a 95 % probability that the value of the measurand is less than the limit *L,* then *g* is chosen, so that for a measured value of $L - g$, the probability that the value of the measurand lies below the limit is 95 % – see Figure 3b.

In most cases the size of the guard band *g* will be a simple multiple of *u*, where *u* is the standard uncertainty. In some cases the decision rule may state the value of the multiple to be used. In general, the acceptance limit will also depend upon the value of the probability, *P,* required, and the knowledge about the distribution of the values of the measurand. In some cases, *g* might be a more complex function of *u*. Some typical cases are described in Annex A.

6 Specifying an acceptable value for standard uncertainty

The larger the value of the standard uncertainty, *u,* and the closer the measured value is to the limit, the larger is the proportion of the samples that will be judged incorrectly. For example, in the case of assessing compliance against an upper limit, when no guard band has been set; if the measured value is more than 3*u* below the limit, then the risk of making a wrong decision is very small (i.e. about 0.1 %). The risk increases as the measured value approaches the limit: it is 2.3 % when the measured value is 2*u* apart from the limit, and 50 % when it coincides with the limit itself. The smaller the value of *u*, in general, the higher will be the cost of measurement. Thus, ideally *u* should be chosen to minimise the cost of measurement plus the cost of the wrong decision.

In some analytical fields, the target (i.e. the maximum) measurement uncertainty [13] is defined together with the maximum and/or permissible limit(s) for the measurand. The Eurachem/CITAC guide on setting and using the target measurement uncertainty [14] suggests how target uncertainty can be defined when it is not regulated or defined by the customer. The target uncertainty can be defined from the width of the permissible interval for the measurand (Section 5.1.2 of the guide), or from a defined acceptance limit for the measured concentration based on an acceptable consumer or producer risk of a wrong decision (Section 5.1.4 of the guide).

NOTE: This page is intentionally empty.

7 Recommendations

In order to decide whether or not to accept/reject an item there has to be:

a) a specification giving the upper and/or lower permitted limits of the characteristics (measurands) being controlled;

b) a measurement uncertainty*; and

c) a decision rule that describes how the measurement uncertainty will be taken into account with regard to accepting or rejecting an item according to its specification and the result of a measurement.

The decision rule should have a well-documented method of unambiguously determining the size of the acceptance and rejection zones, ideally including the minimum acceptable level of the probability that the measurand lies within the specification limits. It should also give the procedure for dealing with, e.g. repeated measurements and outliers (see Section 4.1).

A decision rule can set the size of the acceptance or rejection zone by means of an appropriate guard band. The size of the guard band is calculated from using the knowledge of the measurement uncertainty and the minimum acceptable level of the probability that the measurand lies within the specification limits. For the common cases where the uncertainty is approximately constant or where the assumed error distribution is symmetric with a standard deviation proportional to the "true" value, the uncertainty at the limit can be used to calculate the guard band. This is described in Cases 1-3 in Annex A.

In addition, a reference to the decision rules used should be included when reporting on compliance.

* Including contributions from the sampling process, when the measurand is defined in terms of the sampling target, e.g. a production batch instead of just the laboratory sample.

NOTE: This page is intentionally empty.

Annex A – Determining the size of the guard band and acceptance limit

The size of the guard band g and hence the acceptance limit is chosen to meet the requirements of the decision rule. It depends upon:

1. the value of the uncertainty;
2. the minimum acceptable level of the probability P that the value of the measurand lies within the specification limits (or, equivalently, the maximum acceptable probability that the value of the measurand is not within the specification limit); and
3. the knowledge available about the distribution of the values of the measurand.

Where the relative standard uncertainty is less than about 15 % to 20 %, the distribution can be assumed to be normal [15]. The size of g will then be equal to ku, as in Cases 1a and 1b below. If the effective degrees of freedom are known the value of k will be taken from the t-distribution as in Case 2. In other cases, where it is still known that the value of the measurand is greater than zero, but the relative standard uncertainty is greater than 20 %, the normal distribution may not be appropriate. The size of g is then determined from the shape of the distribution and the desired value of P, as in Cases 3 and 4. There are a number of possible distributions that could be used, e.g. lognormal [15], beta [16] and gamma [9], which give comparable results for relative uncertainties as large as 50 %. Further guidance for assigning a distribution on the basis of the available knowledge can be found in JCGM 101:2008 [17]. When the model equation for calculating the value of the quantity consists of multiplication/division of positive quantities then there are good grounds for utilising the lognormal distribution as described in Case 4 [15].

Case 1a – Standard uncertainty available

In this case, for a relative standard uncertainty of less than 20 %, the size of the guard band will be ku and the value of k will either be specified in the decision rule or will be derived from the probability distribution of the values attributed to the measurand, which is usually assumed to be normal. The basis for making this assumption and the conditions under which it might be appropriate are given in Annex G of the GUM [5]. The assumption is based on the use of the *central limit theorem* and section G 2.3 points out that "…. if the combined standard uncertainty u is not dominated by a standard uncertainty component obtained from Type A evaluation based on just a few observations, or by a standard uncertainty component obtained from a Type B evaluation based on a rectangular distribution, a reasonable first approximation to calculating the expanded uncertainty U that provides an interval with a level of confidence P is to use, for k, the value from the normal distribution".

In many cases, current practice is to use $k = 2$. On the assumption that the distribution is approximately normal, this gives an approximately 95 % level of confidence that, for an observed value x, the value of the measurand lies in the interval $x \pm 2u$. On this basis, the probability that the value of the measurand is less than $x + 2u$ is approximately 97.7 %. In the commonly encountered case of requiring proof of compliance with an upper limit, as shown in Figure 3, taking $k = 2$ and requiring proof of clear non-compliance is equivalent to setting a guard band $g = 2u$. If the observed value exceeds the limit by more than g then the value of the measurand is above the limit with at least 97.7 % confidence. This will therefore result in fewer incorrect non-compliance decisions than decisions based on one-tailed significance tests at 95 % confidence (i.e. with $k = 1.64$*). If it is important to implement decisions at other levels of confidence, then a value of k obtained from tables or statistical software for the appropriate level of confidence can be used.

However, in the GUM [5], section G 1.2, it is pointed out that since the value of U is at best only approximate, it is normally unwise to try to distinguish between closely similar levels of confidence

* The k value is 1.64 with 2 significant figures since the value is 1.6449 with 5 significant digits.

(say a 94 % and a 96 % level of confidence). In addition, the GUM indicates that obtaining intervals with levels of confidence of 99 % or greater is especially difficult.

For a relative standard uncertainty, *u,* greater than 20 % the lognormal distribution could be considered [15]. This is described in Case 4, below, and in Example 3 in Annex B.

Case 1b – Expanded uncertainty available, with a stated value of *k*

Divide *U* by the given value of *k* (normally 2) and determine the value of the guard band using the revised value of *k* appropriate to the application as in Case 1a.

Case 2 – Standard uncertainty and effective degrees of freedom available

In this case it is current practice to assume that the values that could be attributed to the measurand follow a *t*-distribution with known degrees of freedom and using the one-tailed upper 95 % quantile for the *t*-distribution for the coverage factor *k.* The size of the guard band will be *ku* and an example of compliance with an upper limit is shown in Figure 3.

The *t*-distribution and degrees of freedom are discussed in more detail in GUM, sections G3 & G4. Alternative approaches, which avoid the problems with the use of effective number of degrees of freedom, have been proposed by Williams [18] and by Kacker and Jones [19].

Case 3 – Individual components and distributions available

This case is dealt with in Section G 1.4 of the GUM [5]. This states that if the probability distributions of the input variables are known and the value of the measurand is linearly related to these input quantities, then the probability distribution of the values attributed to the measurand can be calculated by convolution of these distributions. This can be done using the Monte Carlo method for propagation of probability distributions and the resulting distribution used to calculate the required confidence interval. For routine use of the convolution method, the Monte Carlo distribution [20] should be compared with other known distributions and in many cases it is likely to show that the lognormal distribution gives a good fit. It is worth mentioning that the Monte Carlo method [17] is applicable to every measurement model linking the measurand to a set of input quantities.

Case 4 – Asymmetric distributions

The case where an input quantity is distributed asymmetrically is covered in Section G 5.3 of the GUM [5]. It points out that “this does not affect the calculation of *u* but may affect the calculation of *U*”.

In general terms, there are three important situations where asymmetric confidence intervals are necessary for taking a decision:

a. When the (assumed) distribution of the measurand *x* is inherently asymmetric (such as the Poisson distribution with low degrees of freedom);
b. When the relative standard uncertainty is 1) larger than 20 % and 2) is constant;
c. When the measured response *x* is close to a physical constraint (e.g. observed concentrations close to zero);

For a and b, when the data used to show that the distribution is asymmetric is available, confidence interval limits can be calculated, as in Case 3 above. The first situation, a, is known from, e.g. radioactivity measurements with a small number of detected events.

The third situation, c, is known from measurements close to a limit of detection or quantification, or when the definition of a variable is limited to a specific interval. In this case it may be necessary to use a truncated distribution (see Ref. [20], Appendix F).

For many analytical measurements, the value of the measurand is known to be positive and the model equation consists of the product or ratio of positive quantities, then for the situations b and c the lognormal is often an appropriate distribution to use.

Assuming that the distribution of the values attributable to the measurand is lognormal the acceptance limits can be calculated using the expanded uncertainty factor ${}^{F}U$ [3, 21]:

$$ {}^{F}U = \exp\,(ks_G) \qquad \textbf{Equation 1} $$

where s_G is the standard deviation in $\log_e$ space (natural logarithms). For u_{rel} less than 0.5 (50 %), $s_G \approx u_{\text{rel}}$ [15] and the uncertainty factor can be calculated as:

$$ {}^{F}U \approx \exp\,(ku_{\text{rel}}) \qquad \textbf{Equation 2} $$

where the coverage factor k is the upper quantile of the standard normal distribution at the desired level of confidence.

The upper acceptance limit for high confidence of correct rejection is then:

$$ L_{\text{u}} \times {}^{F}U \qquad \textbf{Equation 3} $$

and the upper acceptance limit for high confidence of correct acceptance:

$$ L_{\text{u}} / {}^{F}U \qquad \textbf{Equation 4} $$

The guard band for an upper limit for high confidence of correct rejection can then be calculated:

$$ g = L_u \times {}^{F}U - L_u \qquad \textbf{Equation 5} $$

Compared with the normal distribution the size of the guard band for an upper limit for correct rejection will increase (see Figure 3a) and for correct acceptance will decrease (see Figure 3b). This is due to the asymmetry of the lognormal distribution. As an example, the guard bands calculated for a normal and lognormal distribution with an upper limit L of 100, k = 1.64 and u_{rel} of 0.3 and 0.5, respectively, are given in Table 1. The relative standard uncertainties of 0.3 and 0.5 correspond to expanded relative uncertainties of 60 % and 100 % respectively.

Table 1 — Acceptance limits for an upper limit assuming a normal and a lognormal distribution for high relative uncertainty

Assumed distribution	Relative standard uncertainty, u_{rel}	Upper limit	Acceptance limits	
			Correct acceptance	Correct rejection
Normal	0.3	100	51	149
Lognormal	0.3	100	61	164
Normal	0.5	100	18	182
Lognormal	0.5	100	44	227

In terms of u_{rel}, the acceptance limits for the normal distribution are:

$$L(1 + ku_{\mathrm{rel}}) \text{ and } L(1 - ku_{\mathrm{rel}})$$

and for the lognormal are:

$$L\,(\exp(ku_{\mathrm{rel}})) \text{ and } L(\exp(-\,ku_{\mathrm{rel}})).$$

For the lognormal distribution the equations can be expanded (using the usual expansion for $\exp(x)$) to:

$$L\left(1 + ku_{rel} + \frac{(ku_{\mathrm{rel}})^2}{2} + \cdots\right) \text{ and } L\left(1 - ku_{\mathrm{rel}} + \frac{(-\,ku_{rel})^2}{2} + \cdots\right),$$

where "…" denotes higher terms in the expansion. When the terms above ku_{rel} are significant the use of the lognormal distribution should be considered [15]. At $u_{\mathrm{rel}} = 20\ \%$ and $k = 1.64$, the increase of the factor for calculating the acceptance limit for correct rejection will be about 5 % compared with using $(1 + ku_{\mathrm{rel}})$.

An example of the use of the lognormal distribution is given in Annex B, example 3.

Annex B – Examples

Example 1 - Implementation of a decision rule covered by Case 1b in Annex A

Case 1b in Annex A deals with the situation when the expanded uncertainty is available, together with a stated value of *k*. The nickel mass fraction for a type of stainless steel must be in the range from 16.0 % to 18.0 % Ni.

Measurand	Mass fraction of nickel, Ni in a batch of steel delivered to a customer.
Uncertainty	The absolute expanded uncertainty, *U,* is 0.2 % Ni, *k* = 2 (95 %). Standard uncertainty, *u* = 0.1 % Ni. This uncertainty includes both sampling uncertainty for the batch and analytical uncertainty.
Specification	The specification zone is from the lower limit 16.0 % Ni to the upper limit 18.0 % Ni.
Decision rule *High confidence of correct acceptance*	*The acceptance zone is the mass fraction interval where it can be decided with a confidence level of not less than 95 % (α=0.05) that the batch has a mass fraction above the lower limit and below the upper limit.*
Distribution	The distribution of the values of the measurand is assumed to be Normal.
Guard band	Each guard band is calculated as $1.64u \approx 0.17$ % with *k* value 1.64 from the one-tailed 95 % upper quantile for the normal distribution.
Acceptance zone	16.2 % Ni to 17.8 % Ni, after rounding to one decimal place.
Measured value	16.1 % Ni

Figure 4 shows the positions of the acceptance and rejection zones for a specification with upper and lower limits with high confidence of correct acceptance. The measured value, 16.1 % Ni is below the lower acceptance limit of 16.2 %, i.e. it is in the rejection zone. **The batch is non-compliant**.

NOTE
If the decision rule stated simple acceptance, the acceptance zone would be 16.0 % to 18.0 % and the batch would be compliant.

Example 2 - Implementation of a decision rule covered by Case 2 in Annex A

Case 2 in Annex A deals with the situation when the standard uncertainty, u, and effective degrees of freedom are available. In a production batch the concentration of an analyte should be below 200 ng/g.

Measurand	Mass fraction of an analyte in a production batch delivered to a client.
Uncertainty	The absolute standard uncertainty u = 2.2 ng/g. The uncertainty includes components arising from sampling. The dominant contribution to uncertainty is based on 9 measurements, i.e. 8 degrees of freedom and it can be assumed that the values attributable to the measurand follow the *t*-distribution.
Specification	Upper permitted limit $L_u$is 200 ng/g.
Decision rule *High confidence of correct rejection*	*The batch will be considered to be non-compliant if the probability of the value of the mass fraction being greater than 200 ng/g exceeds 95%.*
Distribution	The distribution of the values of the measurand is assumed to be Normal.
Guard band	The guard band is calculated with a k value of 1.86 from the one-tailed upper 95 % quantile for the *t*-distribution with 8 degrees of freedom to be: $ku = 1.86 \times 2.2 = 4.1$ ng/g.
Acceptance limit	204.1 ng/g
Measured value	203.7 ng/g

Figure 3a shows the positions of the acceptance and rejection zones for a specification with an upper limit with high confidence of correct rejection. The measured value, 203.7 ng/g is under the acceptance limit of 204.1 ng/g, i.e. it is in the acceptance zone. **The batch is compliant**.

NOTE
If the decision rule stated simple acceptance, the acceptance limit would be equal to the permitted limit of 200 ng/g and the batch would be non-compliant.

Example 3 - Implementation of a decision rule covered by Case 4 in Annex A

Case 4 in Annex A deals with asymmetric distributions. In this case the asymmetry is due to high relative uncertainty in the analysis for the control of a banned substance and the distribution for the uncertainty is approximately lognormal.

Measurand	Mass fraction of a banned substance in a sample.
Uncertainty	The relative standard uncertainty u_{rel} is 35 %.
Specification	Upper permitted limit L_{u} is 2 ng/g.
Decision rule *High confidence of correct rejection*	*The concentration of the banned substance will be deemed to be above the limit if the probability of the value of the concentration being greater than the limit is ≥ 95 %.*
Distribution	The distribution of the values of the measurand is assumed to be lognormal.
Guard band	For a lognormal distribution the guard band can be calculated using the uncertainty factor ${}^{F}U$ from Equation 2, ${}^{F}U \approx \exp(ku_{\text{rel}})$; k is 1.64, from the one-tailed upper 95 % quantile of the normal distribution and $u_{\text{rel}} = 0.35$, giving ${}^{F}U \approx \exp(1.64u_{\text{rel}}) = 1.78$. The guard band g for *correct rejection* can then be calculated as: $g = L_u \times {}^{F}U - L_u = 1.6$ ng/g.
Acceptance limit	3.6 ng/g
Measured value	3.3 ng/g

Figure 3a shows the positions of the acceptance and rejection zones for a specification with an upper limit with high confidence of correct rejection. The measured value, 3.3 ng/g is below the acceptance limit of 3.6 ng/g, i.e. it in the acceptance zone. **The sample is compliant**.

NOTE
The assumption of type of distribution is crucial. If a normal distribution were assumed in this case, the acceptance limit would be lower, at 3.2 ng/g, and the sample would not be compliant, $L + g = L(1 + ku) = 2(1 + 1.64 \times 0.35) = 3.2$. A comparison between the lognormal and the normal probability distributions is shown in Figure 5.

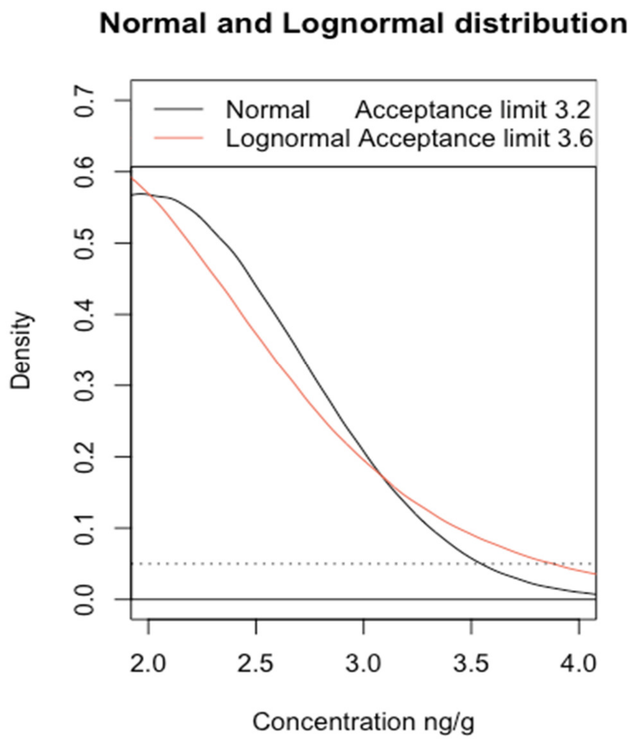

Figure 5 ─ Right side of probability distributions for a limit value of 2 ng/g with a relative standard uncertainty of 35 % for normal and lognormal distributions showing the difference in upper tails resulting in different acceptance limits. Horizontal line at density 0.05 to aid visual comparison

Annex C – Producer and consumer risk

Introduction

The body of this Guide is concerned with the use of decision rules given a result and an associated expanded uncertainty. Some guidance, for example ILAC G8 [6], includes recommendations for setting decision rules intended to control producer and/or consumer risk. This Annex provides a brief explanation of producer and consumer risk, and of "specific" and "global" risk.

Producer and consumer risk

Producer and consumer risk are concepts arising from manufacturing process management, though they apply well to many compliance situations and are used, e.g. in "acceptance sampling" of products. In a manufacturing environment, the "producer risk" is just the probability of incorrect rejection of acceptable products – so-called because this results in an unnecessary cost to the producer. Similarly, "consumer risk" is the probability of incorrect acceptance of non-compliant products; the chance that a consumer might receive a faulty product that has passed inspection.

The ideas are illustrated in Figure 6. The top curve – the "process distribution" – is the distribution of values associated with products produced by a manufacturing process. The values are for some important measured characteristic; for example, drug dosage in a pharmaceutical, packaged weight of a food product, or alcohol concentration in a beverage. L_l and L_u are permitted lower and upper limits for the characteristic; for simplicity, the figure assumes that these are also set as acceptance limits for inspection – that is, there is no guard band. A product between L_l and L_u is compliant; a product outside is non-compliant. The value at A in the figure is non-compliant. The distribution associated with the non-compliant value A is the distribution of measurement results that might be observed on testing products with this value of the characteristic. A proportion (shaded) of these results fall inside the acceptance limits; this proportion is the false acceptance rate for a product at A. It is an example of a consumer risk; the probability of acceptance of noncompliant product.

The value at B in Figure 6 illustrates a producer risk. The value B is within permitted limits but there is a probability (shaded portion of the distribution for results at value B) of results falling outside acceptance limits; this proportion is a producer risk.

> NOTE The description above is based on the traditional model for producer and consumer risks, which assumes a process that generates products with true values such as A and B, and a subsequent measurement error distribution which leads to a distribution of measurement results used for decisions. This is a theoretical model. More recent views start with measured values, uncertainties and limited information about the process, and consider the conclusions and probabilities that can be drawn from these. This later view is considered briefly below.

Specific and global risks

The proportion of apparently compliant measurement results in Figure 6 shown as a shaded proportion of results associated with the value A is specific to products with that value of the characteristic. It is an example of "specific risk"; the probability of an incorrect compliance decision associated with product at a particular value. In the case of value A, it is a specific consumer risk for value A. Similarly, the producer risk shown as the shaded portion of the distribution for value B is the specific producer risk for value B.

An important feature of the specific risk is that it depends almost solely on the distribution of measurement results for a given true value of the characteristic. From the perspective of a testing laboratory with a measurement result, the laboratory's estimate of the specific risk depends on the measured value and the measurement uncertainty. For any given value for the test item (that is, the product), the specific risk is smaller when the uncertainty is smaller.

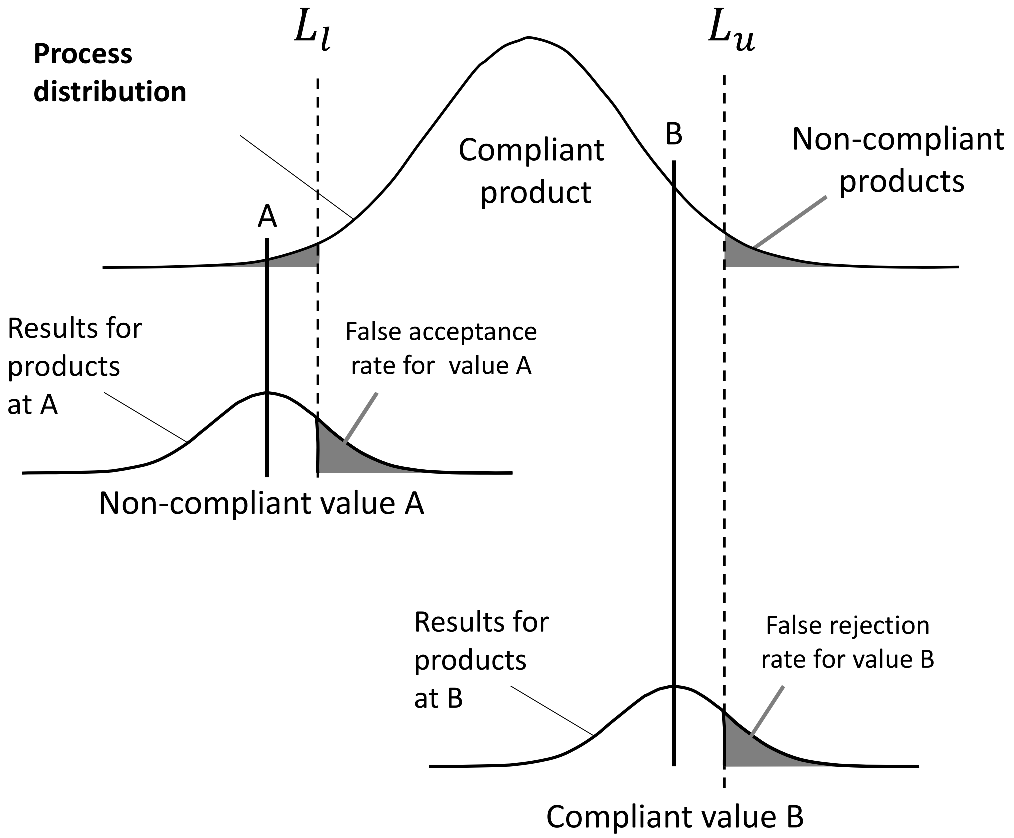

Figure 6 ─ Producer and consumer risk. The figure shows the distribution (top) of values of a measurand of interest generated by a manufacturing or other process, with permissible range between limits L_l and L_u, together with distributions for measured values from items at A and B. See text for further details

Specific risk, however, does not describe the overall probability of incorrect decisions of each kind, because test items – products – with different values each have their own specific risk. There is therefore a second important probability, termed the "global risk". The global risk is the probability of incorrect decisions taken over the whole production distribution. For the case of consumer risk, the global consumer risk is the combined probability of incorrect acceptance decisions; the combination of the specific risks for all possible values for non-compliant product, weighted by their frequency of occurrence. Similarly, the global producer risk arises from the combination of all the specific producer risks at all the different values between L_l and L_u.

Note: Global risk is calculated as the sum of all specific risks at each possible value, multiplied by their probability of occurrence. For a continuous distribution like those shown in Figure 6, the probability of occurrence is replaced by the height of the curve describing the process distribution (the density), and the sum becomes an integration over both process and measurement distributions. The mathematical detail is given in, for example, JCGM 106 [9].

An important difference between specific and global risks is that the global risk depends strongly on the process distribution, whereas the specific risk does not. For example, for a hypothetical manufacturing process that only generates compliant product, the global consumer risk can only be zero because there is no possibility of an inspection result passing non-compliant product. Similarly, for a very poor process with a high probability of generating non-compliant material, the global consumer risk will be comparatively high.

For the testing laboratory, however, the distribution of values produced by a process is often unknown. It is for this reason that a testing laboratory will find it easiest to rely on specific risk rather than global risk. In addition, if the specific risks are kept small – particularly by keeping uncertainties small – the global risks will also be kept small. ILAC G8 [6] accordingly advises that, where there is no other basis for setting decision rules, the decision rule should be set to keep the specific consumer risk low.

Specific risk for a measurement result

It was noted above that the general description in Figure 6 is based on the theoretical distributions for values arising from a production process and for measurement results. The former are "true values" for products leaving a process; the latter is the distribution of observed values given (true) values for test items. In practice, a measurement laboratory only has measurement results with uncertainties, and may sometimes have information on the values expected from a process, whether natural or commercial. The measurement laboratory can therefore only estimate specific and global risks from the information they have.

When the information about the production process is only weakly informative or when the uncertainty is small compared to the width of the process distribution, specific risk can be estimated adequately from the measurement uncertainty and its associated distribution. The relevant risk is just the proportion of the uncertainty distribution beyond the relevant permitted value. This is identical to the shaded regions for values A and B in Figure 6.

When there is substantial information on the process distribution, however, the specific risk for a particular test item can be calculated to take account of the prior probability that the tested item is compliant, based on the process distribution. This can change the estimated specific risks, sometimes appreciably. A full treatment is not within the scope of the present Guide; full details are, however, given in JCGM 106 [9].

NOTE: This page is intentionally empty.

Annex D – Definitions

The following definitions are based on definitions in ASME B89.7.3.1-2001 [7], VIM [4], GUM [5], ILAC G8 [6] or ISO/IEC 17025 [8].

measurand: particular quantity subject to measurement

expanded uncertainty: quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand

measurement result: set of quantity values being attributed to a measurand together with any other available relevant information

NOTE: A measurement result is generally expressed as a measured value and an uncertainty interval.

decision rule: rule that describes how measurement uncertainty is accounted for when stating conformity with a specified requirement

specification limit (tolerance limit): specified upper or lower bound of permissible values of a property

specification zone (tolerance zone): interval of permissible values of a property

acceptance limit: specified upper or lower bound of permissible measured quantity values

simple acceptance: a decision rule in which the acceptance limit is the same as the specification limit

acceptance zone (acceptance interval): the set of values of a characteristic, for a specified measurement process and decision rule, that results in product acceptance when a measurement result is within this zone

rejection zone (rejection interval): the set of values of a characteristic, for a specified measurement process and decision rule, that results in product rejection when a measurement result is within this zone

guard band: interval between a specification limit and a corresponding acceptance limit

NOTE: This page is intentionally empty.

Bibliography

For updates to referenced guidance and for additional reading please refer to the Eurachem Reading List on the Eurachem website, www.eurachem.org.

1. I. Kuselman, F. Pennecchi, R. Bettencourt da Silva, D.B. Hibbert, IUPAC/CITAC Guide: Evaluation of risks of false decisions in conformity assessment of a multicomponent material or object due to measurement uncertainty, (IUPAC Technical Report), Pure Appl. Chem., (2020). Available from doi.org/10.1515/pac-2019-0906.
2. F. Pennecchi, M. G. Cox, P. Harris, A. M. H. van der Veen, S. L. R. Ellison, Euramet project EMUE-D2-1-Multicomponent Materials (2020). Available from www.euramet.org.
3. M. H. Ramsey and S. L. R. Ellison (eds.) Eurachem/EUROLAB/CITAC/Nordtest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches (2nd ed. 2019). ISBN 978-0-948926-35-8. Available from www.eurachem.org.
4. JCGM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Available from www.bipm.org.
5. JCGM 100:2008, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM). Available from www.bipm.org.
6. ILAC G8:2019, Guidelines on Decision Rules and Statements of Conformity. Available from www.ilac.org.
7. ASME B89.7.3.1-2001, Guidelines for decision rules: considering measurement uncertainty in determining conformance with specifications (asme.org).
8. ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories, ISO, Geneva, (2017).
9. JCGM 106:2012, Evaluation of measurement data – The role of measurement uncertainty in conformity assessment. Available from www.bipm.org.
10. Á. Silva Ribeiro and M. Golze, EUROLAB Technical Report 1/2017: Decision rules applied to conformity assessment, Available from www.eurolab.org.
11. WADA Technical Document – TD2019DL, Decision limits for the confirmatory quantification of threshold substances (2019). Available from www.wada-ama.org.
12. ISO 10576:2003 Statistical methods — Guidelines for the evaluation of conformity with specified requirements — Part 1: General principles, ISO, Geneva, (2003).
13. Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (2002/657/EC).
14. R. Bettencourt da Silva and A. Williams (eds), Eurachem/CITAC Guide: Setting and using target uncertainty in chemical measurement, (1st ed. 2015). Available from www.eurachem.org.
15. A. Williams, Calculations of the expanded uncertainty for large uncertainties using the lognormal distribution, Accred. Qual. Assur., **25**, 335-338 (2020).
16. A. M. H. van der Veen and G. Nieuwenkamp, Revision of ISO 19229 to support the certification of calibration gases for purity, Accred. Qual. Assur., **24**, 375–380 (2019).

17. JCGM 101:2008, Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method. Available from www.bipm.org.
18. A. Williams, An alternative to the effective number of degrees of freedom, Comparability and Reliability in Chemical Measurement, Accred. Qual. Assur., **4**, 14 – 17 (1999).
19. R. Kacker and A. Jones, On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent, Metrologia, **40**, 235-248 (2003).
20. S. L. R. Ellison and A. Williams (eds), Eurachem/CITAC Guide: Quantifying uncertainty in analytical measurement, (3rd ed. 2012), ISBN 0 948926 15 5. Available from www.eurachem.org.
21. M. H. Ramsey S. L. R. Ellison S.L.R Uncertainty Factor: an alternative way to express measurement uncertainty in chemical measurement. Accred. Qual. Assur., **20**, 153-155 (2015).

ISBN 978-0-948926-38-9