CODEX ALIMENTARIUS COMMISSION

Food and Agriculture Organization of the United Nations

World Health Organization

Viale delle Terme di Caracalla, 00153 Rome, Italy - Tel: (+39) 06 57051 - Fax: (+39) 06 5705 4593 - E-mail: codex@fao.org - www.codexalimentarius.org

Agenda Item 6 **CX/MAS 18/39/6**

April 2018

JOINT FAO/WHO FOOD STANDARDS PROGRAMME
CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING
39th Session
Budapest, Hungary, 7 – 11 May 2018

PROPOSAL TO AMEND THE GUIDELINES ON MEASUREMENT UNCERTAINTY (CXG 54 – 2004) (Report of the Electronic Working Group led by Germany)

Introduction

At the 38th session of CCMAS (CCMAS38), it was noticed that the proposed revision of the *Guidelines on Measurement Uncertainty* (CXG 54 – 2004) (GL54) would envisage new work for CCMAS. It was agreed to establish an EWG chaired by Germany to work out a clear outline of what the work would entail. This would be outlined in form of a project document, to be discussed at CCMAS39.

The document at hand reflects the purpose and rationale on the amendment of the *Guideline on measurement uncertainty* (CXG 54-2004). It also summarises the work of the EWG after its formation after CCMAS38 in order to identify improvements and propose changes.

Background

As early as CCMAS33 in 2010, much time has been spent discussing amendments to provide a revised draft serving as a basis of possible improvements. The revision of the guideline originates from the requests for more detailed explanations regarding the impact of measurement uncertainty (MU) of particular analytical test samples, sampling procedures, lot assessment and its role in conformity assessment. Some members found it necessary to clarify why MU is important, what kind of influence MU will have on decision-making and its role in conformity assessment of a particular analytical test sample.

A clear assignment will be required, what type of measurement uncertainty the guide will cover.

This should come with practical examples, referring to corresponding international standards. These recommended procedures are necessary for the determination of uncertainty of measurement results, including sub-sampling, sample processing and analysis. The information about the reported expanded measurement uncertainty should comply with the corresponding international standards.

The original intention was to keep the GL 54 as simple as possible. By adding a large amount of texts and examples to substantiate the demand for practical examples for calculating MU in various particular situations would extent the revised guide. That might contradict the original aim.

The specific terms of reference (TOR) of the EWG were arranged at the CCMAS38 session and are as follows:

- Preparation of a project document that indicates which amendments and improvements should be identified and used in GL54 (TOR 1).
- Revision of GL54 considering the identified areas of improvement and technical and other amendments taking into account the need to simplify the content (TOR 2).
- Elaboration of an information document with examples of procedures for estimating measurement uncertainty (TOR 3).

It might be discussed whether GL 54 should be extended for practical use, providing more than general aspects on MU, as this has been done in the *General Guidelines on Sampling* (CXG 50 – 2004) for sampling plans and the *Guidelines on Estimation of Uncertainty of Results* (CXG 59 – 2006) for MU in pesticide analysis,. In any case, it was agreed upon to avoid any kind of overlapping with the *Guidelines on Estimation of Uncertainty of Results* (CXG 59-2006).

Main aspects to be covered:

- An updated GL 54, which is comprehensive, simple to use and understood by Codex commodity committees. It includes a prioritised combination of general and technical improvements.

- Deals with general aspects of MU, which includes the expanded MU (Introduction part to the CXG 54 – 2004).
- Statement of MU and its influence on sampling and the impact on conformity assessment.
- A separate information document sums up recommended procedures for determine uncertainty of measurement results.
- The MU comprises only laboratory samples and solely concerns the uncertainty of results for laboratory test samples, including subsampling.

What the guide will not cover:

- The updated GL 54 does not cover uncertainty, which is derived by sampling.
- No clear explanation is given so far, how acceptance sampling and measurement uncertainty will influence the conformity assessment.
- MU might be an influencing parameter on the decision-making process regarding conformity assessment, but criteria of product assessment are to be harmonized between the trade partners.
- Diversification of national legislations and apparent resultant or potential impediments to international trade are not covered.

Further points for discussion:

- Work is already undertaken by other international organizations in the field of measurement uncertainty. It is observed by the relevant international intergovernmental body(ies), that there is a noticeable lack of advice how these standards on measurement uncertainty are linked to sampling and conformity assessment procedures. It needs to be discussed, to what extent MU influences the conformity assessment and how comprehensive this topic should be covered by the revised GL 54.

EWG process and discussion

Germany officially invited members and observers to participate in the EWG “Revision of the Guidelines on measurement uncertainty” (CCMAS-GL-mu) via an electronic platform, supplied by the Codex Secretariat. The invitation was send out in September 2017. The EWG has got 52 members. For a list of participants please see Appendix IV.

Early in 2018, three documents were prepared and provided through the electronic platform, asking members for their comments:

TOR 1: Project document

TOR 2: Revision of GL54

TOR 3: Information Document

Response to TOR 1

Germany received comments from 1 Member (Japan), which included 2 technical comments and 2 editorial comments

Summary of main changes:

- It was suggested to remove the description regarding the use of MU in the conformity assessment since conformity assessment is decided by the competent authorities of each country
- In addition, it was emphasised that the GL-54 does not comprise the scope of MU, which arises from sampling. It was proposed to minimise this part.

Response to TOR 2

Comments were received from 10 members (Chile, Iran, Japan, Kazakhstan, Norway, Switzerland, Thailand, Uruguay, Colombia and New Zealand) and included 34 technical comments, 19 general comments and 8 editorial comments.

Summary of main changes:

- Participants wished for readability and an improved structure, including a logical relation in the order of the topics. Regarding the structure of the document, suggestion were made to adopt the

whole format towards ISO standards. Some members expressed their wish to re-locate certain parts of a particular topic towards another location of the document to improve readability.

- Suggestions were also made to remove entire descriptions or minimise particular description, aiming to simplify the document. On the other hand, propositions have been made to provide more and in detailed explanations with examples regarding MU.
- It has been stated, that CCMAS has not reached a complete agreement on MU and its role in conformity assessment. There have been contradictory suggestion to leave out the use of MU in conformity assessment or to provide general advice on MU, its use for the evaluation of conformity and its relationship with sampling.

Response to TOR 3

Two member countries commented on the information document (Colombia and Japan) of which 6 were technical comments, 2 general comments and 5 editorial comments.

Summary of main changes:

- It was suggested that more information on bias and errors (systematic and random) may be added with regard to defining methods in consideration to subsample inhomogeneity and sample preparation variability.
- In the same way, suggestions have been forwarded to include uncertainty contribution of subsampling.
- As it has been mentioned before (see TOR 1 and TOR 2), that participants of the EWG regard conformity assessment to be decided upon by the competent authorities of each country.
- It was proposed to minimise the sampling plans, since uncertainty, which was derived from sampling was not agreed to be considered.

Summary:

The purpose of the proposed new work is to further revise and amend the Guidelines on Measurement Uncertainty (CXG 54 – 2004). It originated from the concern that MU of particular analytical test samples has got an impact on decision making and conformity assessment. This particular topic needs to be addressed regarding the comprehensiveness of the revised GL 54.

It should be mentioned here, that MU deals with laboratory samples and not with the homogeneity of a lot (the guide does not concerns uncertainty, which is derived by sampling). This was agreed upon by the current work assignment (REP17/MAS).

No clear classification is given so far how acceptance sampling and measurement uncertainty will influence the conformity assessment. It would be advisable to discuss a clear scope and to agree upon how extensive this topic should be taken care of within the revised GL 54.

Measurement uncertainty might be an influencing parameter on the decision-making process regarding conformity assessment. The criteria of product assessment are to be harmonized between the trade partners. The revised guide GL 54 was intended to be applicable to all commodities committees and to give national legislative authorities, food producers and traders a helping guide to harmonize their policies on international grounds.

RECOMMENDATIONS

The Committee is invited to:

1. Agree on the proposal for new work to revise GL54 (See project document in Appendix 1;
2. Consider the proposed draft revision of GL 54 (Appendix II) and the proposed draft information document (Appendix III).

APPENDIX I

Project document for New Work to revise the guidelines CAC/GL 54 - 2004 – Measurement Uncertainty

The purpose and the scope of the standard

The purpose of the proposed new work is to revise the Guidelines on Measurement Uncertainty (CXG 54 – 2004) in order to improve and simplify the content.

Relevance and timeliness

The revision of the Guidelines originates from the requests for more detailed explanations regarding the impact of measurement uncertainty of particular analytical test samples regarding sampling procedures, lot assessment and its role in conformity assessment.

It is necessary to clarify why MU is important in its influence on decision-making and its role in conformity assessment of a particular analytical test sample.

The MU deals with laboratory samples and not with the homogeneity of a lot (the guide CXG 54 – 2004 does not concerns uncertainty, which is derived by sampling). A clear assignment will be required what type of measurement uncertainty the guide will cover. No clear classification is given how acceptance sampling and measurement uncertainty will influence the conformity assessment. Measurement uncertainty only concerns the uncertainty of results for laboratory test samples. However, if the measurement uncertainty is significant compared to the uncertainty arising from sampling, it does have an influence on the decision making, whether or not the test sample meets specification. Furthermore, significant measurement uncertainty might have an effect on the sample size as well as on the number of test samples per composite sample of the lot. Since it is essential for the involved authorities for sampling and conformity assessment to understand this relationship, the corresponding amendment of the guidelines are of great importance. Since the involved authorities might not be as familiar with the scope of measurement uncertainty as the laboratories, a general illustrated introduction into that field is recommended.

This should be supported by practical examples, referring to corresponding international standards. These recommended procedures are necessary for determining uncertainty of measurement results, including sub-sampling, sample processing and analysis. The information about the reported expanded measurement uncertainty should comply with the corresponding ISO standards.

Main aspects to be covered

- Introduction part to the CXG 54 – 2004, which covers general aspects on measurement uncertainty, including the expanded measurement uncertainty as well as to emphasise its influence on sampling and its role in conformity assessment.
- An information document will support the revision of CXG 54 -2004
- A prioritised combination of general and technical improvements of an updated CXG 54 – 2004 that is comprehensive, simple to use and understood .

Assessment against the criteria for the establishment of work priorities

General:

Consumer protection from the point of view of health, food safety, ensuring fair practices in the food trade and taking into account the identified needs of developing countries.

Specific:

Criteria applicable to general subjects

- Diversification of national legislations and apparent resultant or potential impediments to international trade: the criteria of product assessment are to be harmonized between the trade partners. Measurement uncertainty might be an influencing parameter on the decision-making process regarding conformity assessment.
- Work is already undertaken by other international organizations in the field of measurement uncertainty. It is suggested by the relevant international intergovernmental body(ies), however there is a noticeable lack of advice how these standards on measurement uncertainty are linked to sampling and conformity assessment procedures.

- Amenability of the subject might be achieved by the appropriate amendments of the existing guidelines CXG 54 – 2004.

Relevance to the Codex strategic objectives

This proposal for new work is within the scope of the Codex Strategic Vision Statement 'To be the preeminent international food standards-setting body to protect the health of consumers and ensure fair practices in the food trade'. The proposed new work item goes in accordance with the Codex 2014–2019 Strategic Plan:

- Strategic goal 1: Establish international food standards that address current and emerging food issues
- Objective 1.1: Establish new and review existing Codex standards, based on priorities of the CAC.
- Activities 1.1.1: Consistently apply decision-making and priority setting criteria across Committees to ensure that the standards and work areas of highest priority are progressed in a timely manner.
- Activities 1.1.2: Strengthen the critical review process to improve standards monitoring.

Information on the relation between the proposal and other existing codex document

A list of issued Codex documents that relate to this proposal are:

- Guidelines on Estimation of Uncertainty of Results (CXG 59-2006). Overlapping of these Guidelines on estimation of uncertainty of results, which has been established by the Codex Committee on Pesticide Residues (CCPR) should be avoided.
- General Guidelines on Sampling (CXG 50 – 2004)

Identification of any requirements for and availability of expert scientific advice

Expert scientific advice might be needed in future

Identification of any need for technical input to the standard from external bodies so that this can be planned for

Technical input from external bodies may also be needed. Up to date drafting and editing according to modern standards are essential to deliver a manageable document.

Overlapping of the Guidelines on estimation of uncertainty of results, which has been established by the Codex Committee on Pesticide Residues (CCPR) should be avoided. Also, there are other important and relevant factors that an updated CXG 54 needs to cover, including the relationship between CXG 54 and other international sources of guidance regarding the subject of measurement uncertainty.

Timeline

Work to start in 2018 after approval by CAC41, with adoption at Step 5 in 2020 and final adoption in 2021.

APPENDIX II

Proposed Draft revised Guidelines on Measurement Uncertainty (CXG 54 – 2004)

Introduction

In 1883 Lord Kelvin once said: “I often say that when you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be” In other words: "to measure is to know" (1).

Many important decisions are based on measurement results of chemical analysis, whether a product is conform to specification, e.g. if water is safe to drink or whether the concentration of a food contaminate (heavy metal ions) is low enough in order to safely consume the food. In some cases, these decisions can implement the destruction of a harvest with financial consequences for the producer. Therefore, it is of upmost importance to ascribe a certain trustfulness to the measurement procedures and its results. Especially if international trade is involved, confidence in data obtained by inspection authorities or testing laboratories is important.

Why is it so important to estimate the measurement uncertainty?

Besides trueness, measurement uncertainty reflects the quality of the result. It enables the user to assess the reliability of the measurement result. This is of upmost importance since testing for regulatory compliance and subsequent decision-making is involved.

Using the guide

Usually laboratories which carrying out chemical analysis do have measurement procedures in place, including effective quality assurance (properly trained staff, maintained equipment, calibration of equipment, reference materials and standards, documentation etc.) This guide assumes these pre-requirements are established.

The guide provides general advice to calculate the measurement uncertainty associated with analytical methods. It should be mentioned here, that the whole topic of uncertainty should be clearly divided into two separate parts: One uncertainty contribution yields from sampling (A), another uncertainty contribution is generated by analysis of the actual sample which usually takes place in a laboratory environment (B). With regard to the decision making process, the final test result will be influenced by both uncertainty contributions. Since the whole topic is a very comprehensive one, this particular guide CAC/GL 54 focuses exclusively on the uncertainty contribution raised by analysis of a particular sample in the laboratory environment (B). For uncertainty arising from sampling please be referred to CXG 50.

Terminology

Measurand, Measurement errors, trueness, precision, accuracy, measurement uncertainty

Measurand

'The measurand is the quantity intended to be measured' (2). This quantity corresponds with a certaint property such as composition, size, temperature, strength….of a measurement object e.g. substance.

Measurement errors, trueness, precision, accuracy

In metrological terms, a measurement error is the difference between a measured value and a reference value (3).

Figure 1 illustrates metrological concepts behind measurement error in form of an archer aiming for the bulls-eye on a target. The bull’s eye on each target represent the “true value” or a “reference value”. Positions 1-4 represent the different archer’s performances with corresponding histograms (gray colour) and probability distributions (red colour) of the positions hit by the archer below. Arrows indicate the position of the bulls-eye, here representing the true value. At position 1, the archer performs with a large scatter, leading to a large random measurement error. The probability is a rectangular and broad scatter. Position 2 reflects a small scatter, hitting consistently the bull’s eye, which leads to a very narrow probability scatter. Example 3 gives a slightly broad performance, which is inclined towards the left side of the target. The archer’s efforts reflecting a systematic error. Consequently, the relevant histogram resembles a normal (Gaussian) distribution which centre position is skewed to the left. On target 4, the archer managed to hit the close proximity of the bull’s eye but some attempts are scattered. The resulting histogram is also a Gaussian distribution with the centre position on the bull’s eye.

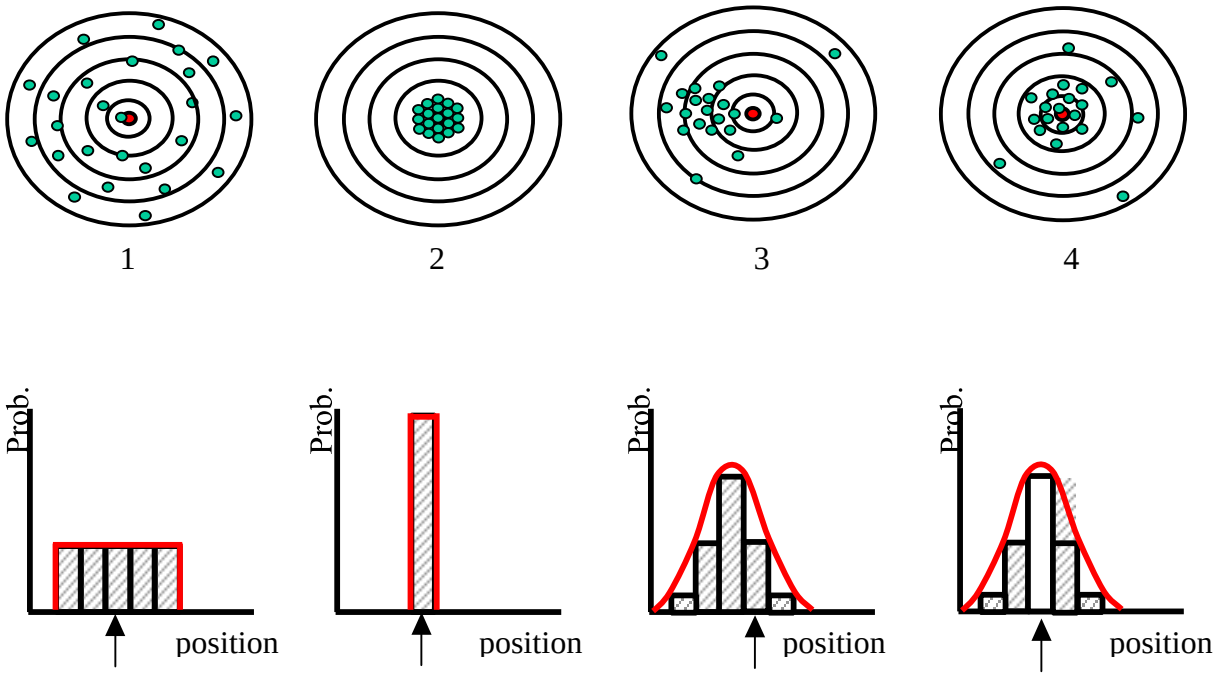

Figure 1: Archer shooting attempts and corresponding histograms (gray colour) and probability distribution (red colour) (3)

(Measurement) trueness is the closeness of agreement between the average value of an infinite number of replicate measured quantity values and a reference quantity value (2). By this definition, archers 1, 2 and 4 each have performed equally well.

(Measurement) precision is the closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions (2). In this sense, the performance of archer 2 has the highest and archer 1 has the lowest precision.

(Measurement) accuracy is the closeness of agreement between a measured quantity value and a true quantity value of a measurand (2). Accuracy requires both trueness and precision. Archer 2 is the most accurate archer. Archer 4 comes closest, with the same trueness but a lower precision score than Archer 2. It is arguable which of the archers 3 or 1 are more accurate: Archer 1 scores best on trueness, Archer 3 scores best in precision (3).

From measurement error to measurement uncertainty

The example of archer 3 reveals a systematic deviation from bull's eye. Performance could be improved by determine the measurement bias (= an estimate of systematic measurement error) and correct for it. Usually this corresponds with good metrological practice. For example, a bias might be associated with a measurement instrument. With the help of a continuous calibration to traceable measurement standards, the instrument bias can be corrected for.

However, time and financial resources do never allow to evaluate and correct for all measurement errors. Consequently, some unquantified errors will remain and inevitable lead to uncertainty about the accuracy of the measurement result, or measurement uncertainty (3).

Measurement uncertainty

The ISO/IEC Guide 99:2007 gives the following definition of measurement uncertainty:

"Non-negative Parameter characterising the dispersion of the quantity values being attributed to a measurand**"** (2).

When scientists make a measurement they generally assume that some exact or **"true value"** exists.

The true value might be for instance an actual calcium ion level in drinking water. But it is almost impossible to measure this “true value” in reality. Measurements per se will never give an exact value. All measurements are imperfect and are subject to systematic errors or random effects like temperature, pressure, humidity fluctuations or even an unsteady judgement of the observer. Even some systematic errors like the drift in between two calibrations are difficult to correct. It should be mentioned here that even the use of a certified reference material comes with a certain measurement uncertainty.

Since the “true value” is impossible to measure, a value is established by using a standardised laboratory method instead. Having included a defined uncertainty range around this measured value, it is very likely that with high probability, the true value is within this range. Half of that uncertainty range is the uncertainty u.

For the purpose of facilitating the abovementioned explanation, Figure 2 illustrates some considerations regarding measurements and its uncertainty (4).

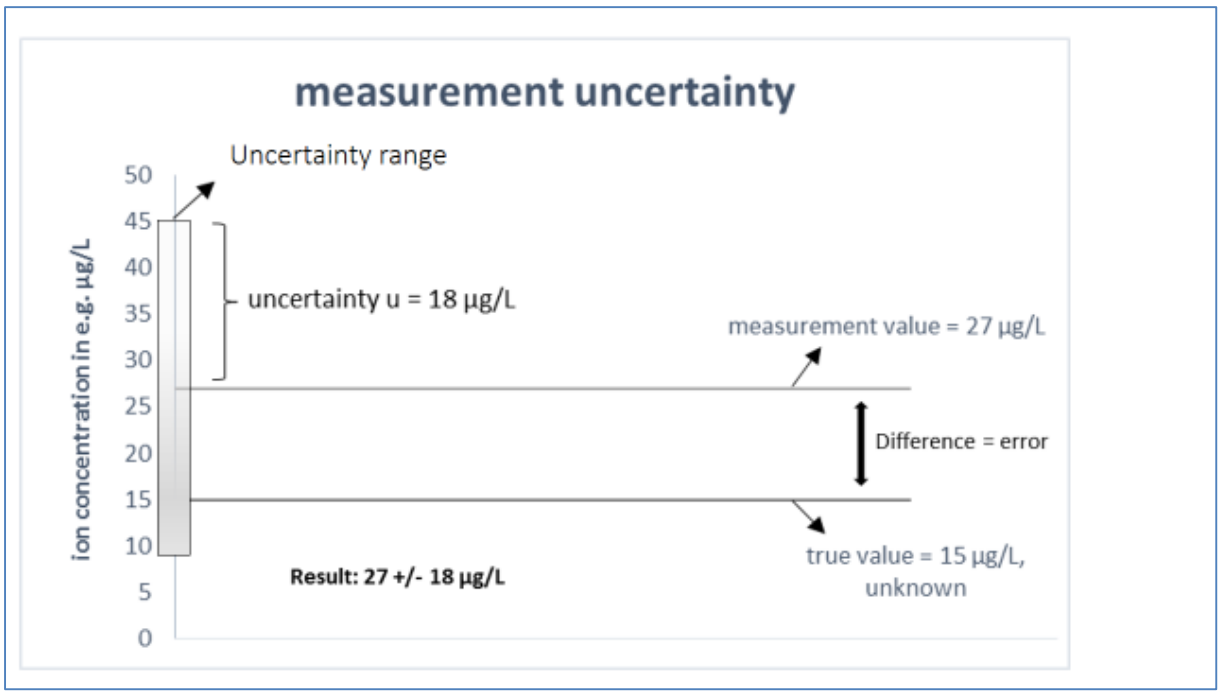

Figure 2: measurement uncertainty

The result of the measurement gives a value of 27 +/- 18 µg/L. There is a high probability, that the true value (which cannot be measured; it remains unknown) of the calcium ion level lies in between the uncertainty range of 45 µg/L and 9 µg/L.

The true value (15 µg/L) in Figure 2 of the calcium ion level is unknown. The estimation of a measurement uncertainty (u = 18 µg/L) can therefore be regarded as the variability (uncertainty range) around the reported result (27 µg/L) of a test sample. Within this particular range, the “true” value can be expected to lie with a reasonable probability.

Therefore, the value of a measurement result needs to be stated together with its uncertainty. In this way, it is easier to judge whether different measured values are comparable or not. Further, when comparing a measurement result with a legally set limit, measurement uncertainty is a helpful aid in order to assess incorrect decisions of compliancy.

Estimating the measurement uncertainty can be done via various approaches. Figure 4 outlines various ways. Typically, the measurement uncertainty is reported as the standard uncertainty multiplied by a coverage factor k = 2, which for a normal distribution corresponds to a coverage probability of approximately 95 %, i.e., the correct value of the measurand is within the range [measured value ± expanded uncertainty] at a confidence level of about 95 % (3).

Notes:

The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

- Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of results of a series of measurements and can be characterised by standard deviations. The other components, which can also be characterised by standard deviations, are evaluated from assumed probability density functions based on experience or other information.
- It is understood that all components of uncertainty, including those arising from systematic effects (bias), such as components associated with corrections and reference standards, contribute to the dispersion.

Uncertainty sources

Many possible sources have been identified, such as incomplete definition of the measurand, subsampling, matrix effects, environmental conditions, uncertainties of lab equipment, reference values to name some (5).

For profound understanding of uncertainty and its sources, Figure 3 gives some explanation and examples regarding uncertainty sources.

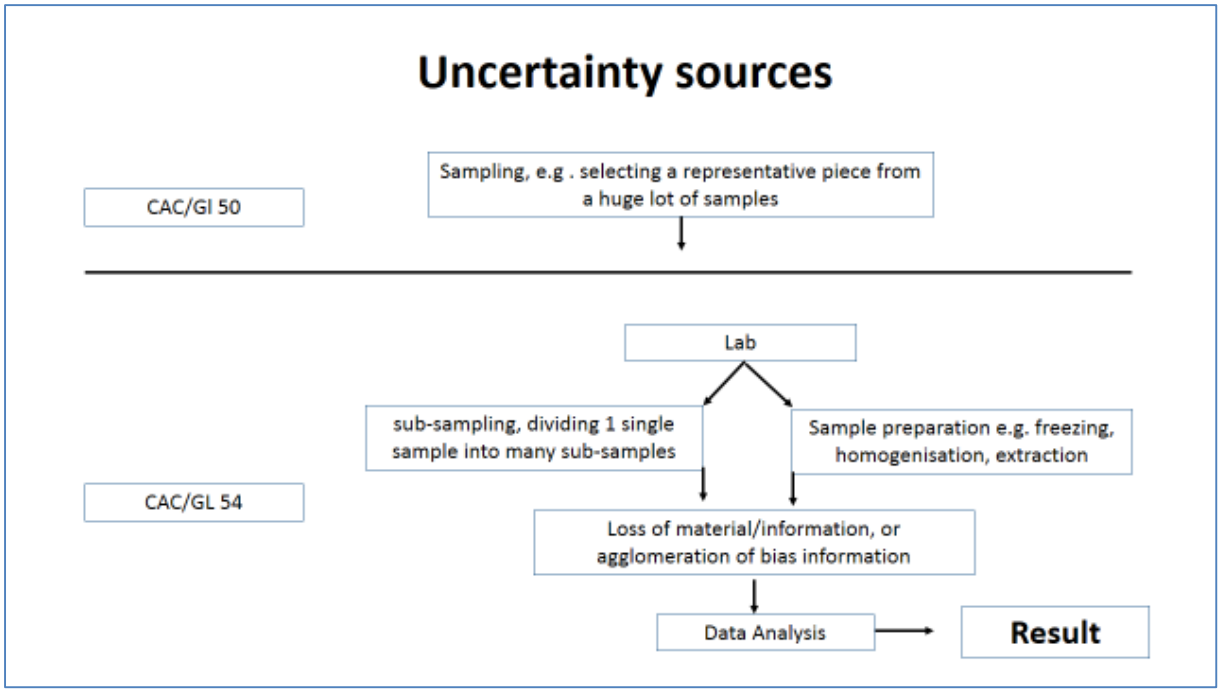

Figure 3: possible uncertainty sources occurring

For more information on the subject, please be referred to EURACHEM guide “quantifying uncertainty in analytical measurement (5).

General requirements for the competence of testing and calibration laboratories are laid down in the ISO/IEC 17025:2005 (6). One of these requirements that Codex has adopted by reference is that testing laboratories should have and should apply procedures for estimating uncertainty of measurement. Information on measurement uncertainty is needed in test reports when it is relevant to the validity or application of the test results, when a customer's instruction so requires, or when the uncertainty affects compliance to a specification limit. The standard also requires that the measurement uncertainty and its level of confidence must, on request, be made available to the user (customer) of the results. (Paragraph 5.4.6).

The Codex Alimentarius Commission has developed *Guidelines for the Assessment of the Competence of Testing Laboratories Involved in the Import and Export Control of Foods* (CXG 27-1997) that require laboratories involved in the import/export of foods to comply with general criteria in ISO/IEC 17025.

Procedures for Estimating Measurement Uncertainty

There are many procedures available for estimating the measurement uncertainty of a result notably those described by ISO (7) and EURACHEM (8). The Codex guidelines do not recommend any particular approach, but it is important that whatever approach is used, the procedure is scientifically credible. No one approach may be said to be better than any other provided the procedure used is

appropriate and credible - i.e. there is no “hierarchy” of the procedures.

In general, procedures are based on a component-by-component (“bottom-up”) approach or on a “top-down” approach using data from collaborative trials, proficiency studies, validation studies or intra-laboratory quality control samples, or a combination of such data (9), (10). Laboratory staff usually know sources of uncertainty because they are professionals and have a good idea about the measurement procedures from experience in using these procedures. In addition, many data is available coming from control charts, Proficiency Tests, calibration data etc.

Figure 4 generalises various approaches to estimate measurement uncertainty.

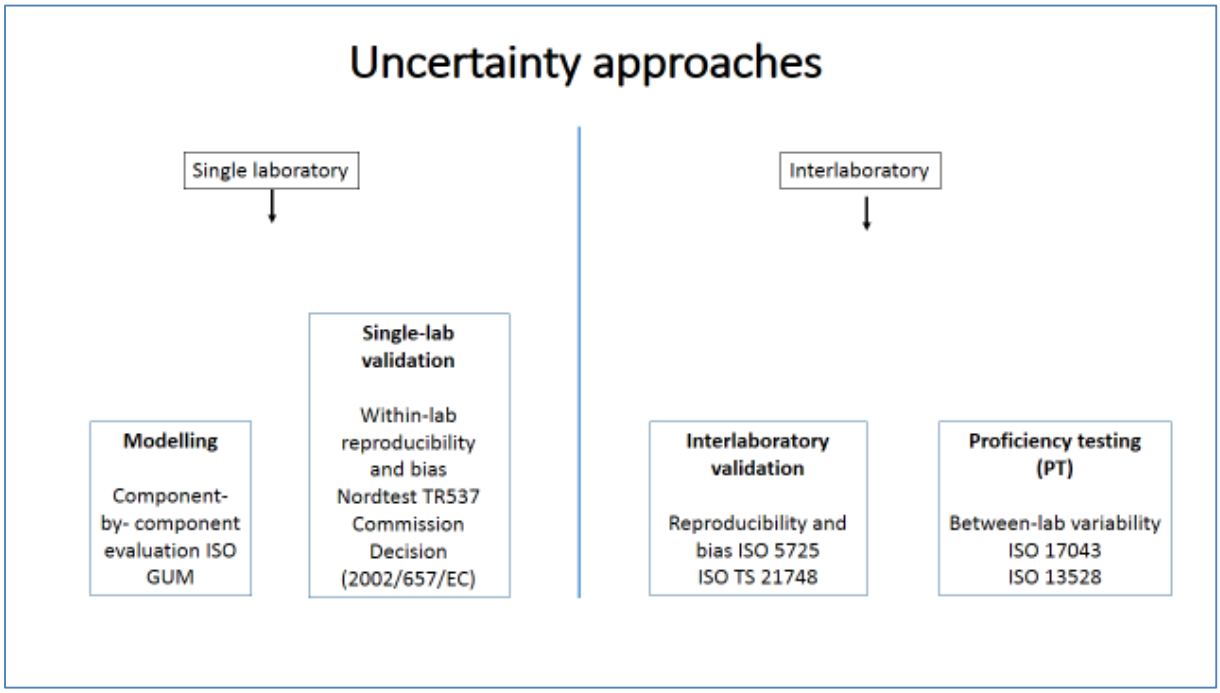

Figure 4: Uncertainty approaches

Regardless of the approach used, the measurand must be clearly defined. Then one needs to consider whether it is a single laboratory approach (Model-based or not Model-based) or if it is an interlaboratory approach using one procedure or not .

The uncertainty calculated by various approaches (Figure 4) can be summarised as follows

- **Modelling** (Classical ISO GUM)
  - Uncertainty of an individual result of a measurement can be obtained, linked to a particular sample
- **Single-lab validation**
  - Typical uncertainty of results obtained using a procedure in the laboratory
- **Interlaboratory validation**
  - Uncertainty of results obtained using the same procedure in different laboratories

In the *Guidelines for the Assessment of the Competence of Testing Laboratories Involved in the Import and Export Control of Foods* (CXG 27-1997) there is a requirement to use validated methods and so it is usually more cost-efficient to use data from the method validation studies rather than using another approach (i.e. the component-by-component approach).It is obligatory for laboratories to check whether parameters were adhered to.

For methods operating within their defined scopes, when the reconciliation stage shows that all the identified sources have been included in the validation study or when the contributions from any remaining sources have been shown to be negligible, then the reproducibility standard deviation s_R, adjusted for concentration if necessary, may be used as the combined standard uncertainty.

It is important that the requirement to estimate measurement uncertainty does not impose any unnecessary additional workloads on laboratories.

When deciding on which procedure is to be used when estimating measurement uncertainty within the Codex context it is important to recognise that Codex has adopted a number of formal quality assurance measures that have to be implemented by control laboratories. In particular, such laboratories should:

- be in compliance with an internationally recognised standard (now with ISO/IEC 17025:2005 Standard); such compliance is aided by the use of internal quality control procedures,
- participate in proficiency testing programmes, and use validated methods.

This section re-emphasises that for the analyst it is important that no unnecessary duplication of existing work is undertaken.

Reported Measurement Uncertainty

The standard measurement uncertainty *u* is the basis for the reported expanded measurement uncertainty *U*. It is obtained by multiplying the standard measurement uncertainty by a coverage factor *k*.

For the level of confidence required (normally 95%), for most purposes it is recommended to set *k=2* (especially for methods which have been validated through collaborative trials). In case that the combined uncertainty is based on only few observations however, *k* should be set equal to the value of Student's t-factor for the so called effective number of degrees of freedom ν_{eff} . (7) , (Annex G.4.1).

Relationship between analytical results and measurement uncertainty

This section attempts to explain the significance of analytical results and their associated measurement uncertainties.

It is important that measurement uncertainty is considered when deciding whether or not a test sample meets the specification. The significance of this can be illustrated by an example shown in the diagram below, which shows the simplest case when decisions are made based on a single test sample. The example shown here is one where the test result is compared against the specification consisting of a maximum level. It illustrates how the concept of measurement uncertainty should be taken into account when interpreting analytical results on a tested sample to allow unambiguous interpretation.

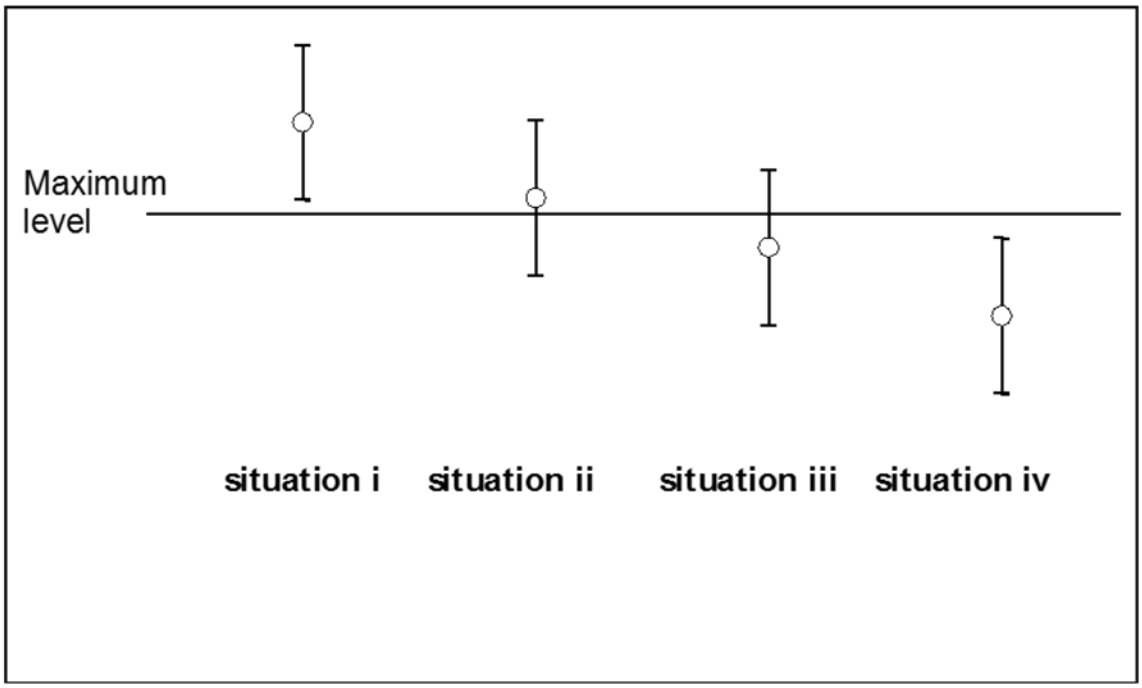

Figure 5: Comparing test results with a Maximum Level taking into account the expanded measurement uncertainty

Situation i

The analytical result minus the expanded measurement uncertainty exceeds the maximum level. The

result indicates that the measured analyte in the test sample is above the specification.

Situation ii

The analytical result exceeds the maximum level by less than the expanded measurement uncertainty.

Situation iii

The analytical result is less than the maximum level by less than the expanded measurement uncertainty.

Situation iv

The analytical result is less than the maximum level by more than the expanded measurement uncertainty.

Note:

Obviously, in the situations ii and iii, the suggested procedure for use of measurement uncertainty in sample assessment can allow acceptance of samples whose true values lie above the maximum level. But the probability of non-compliance is less than the required 95%.

The implications of the situations i to iii in case of testing MRL compliance are extensively discussed in the *Guidelines on estimation of uncertainty of results* (CXG 59-2006).

Relationship between measurement uncertainty, lot conformity assessment and sampling plans

There should be no confusion between the activities of conformity assessment and acceptance sampling. Measurement uncertainty only pertains to the uncertainty of results for laboratory test samples. It does not cover the uncertainty involved with sampling from a lot of product.

Hence, for quantitative estimations on test samples for inspection by variables and inspection by attributes acceptance of a lot is based on the criteria of the corresponding sampling plans i.e. ISO standards.

However, if the measurement uncertainty is **not negligible or dominant** compared to the sampling uncertainty (which is to be proven just by estimation of the measurement uncertainty), for **inspection by attributes** (in case of a quantitative characteristic), it does have influence on the decision whether or not test samples meet the specification i.e. on the verification of the acceptance/rejection number (11).

Example:

A single sampling plan by attributes of AQL = 2,5 % is used to inspect the sodium content of a lot (25 items) of dietary cheese low in sodium. A nonconforming sample increment is one whose sodium content is higher than the maximum sodium content i.e. 120 mg/100g (CXS 53-1981), taking into account the expanded measurement uncertainty.

The decision to be taken according to this sampling plan (11), (Tables 1 and 2A) is to accept the lot if there is no nonconforming sample increment (acceptance number c = 0) in a sample of five increments (n = 5).

The following Figure 6 and Figure 7 compare the situations in two imaginary Laboratories 1 and 2 with equal mean values (points) but with different expanded measurement uncertainties (bars).

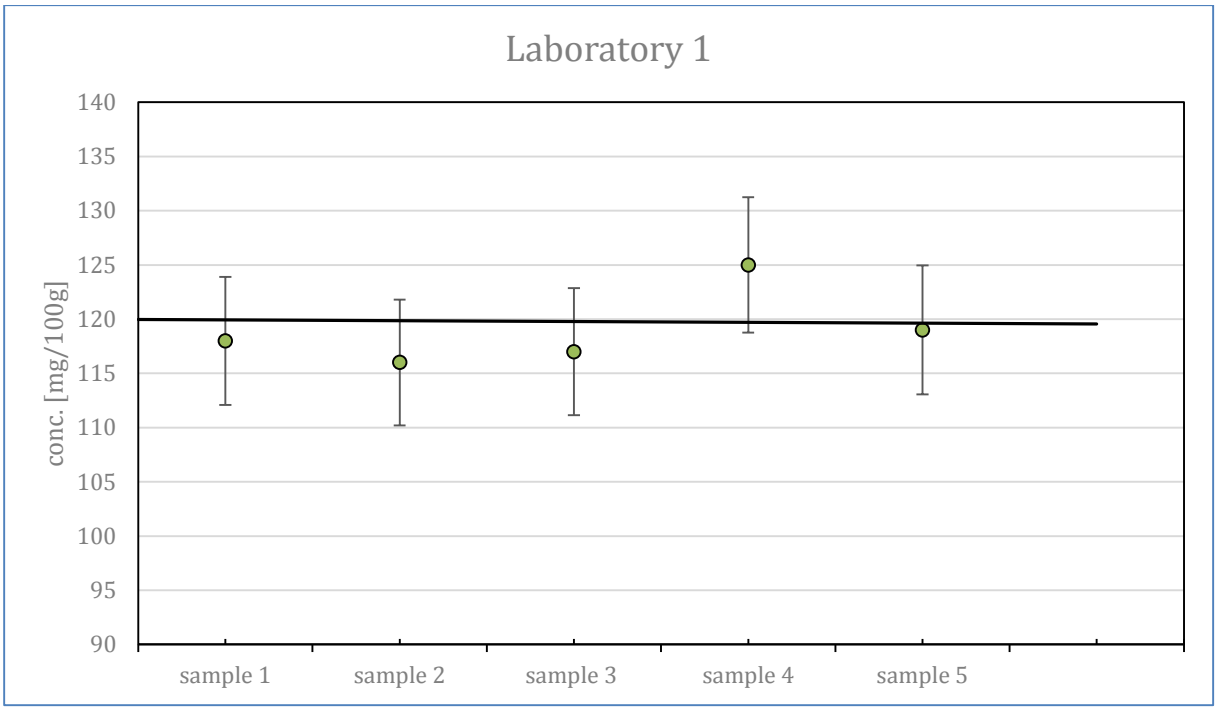

Figure 6: Laboratory 1 results

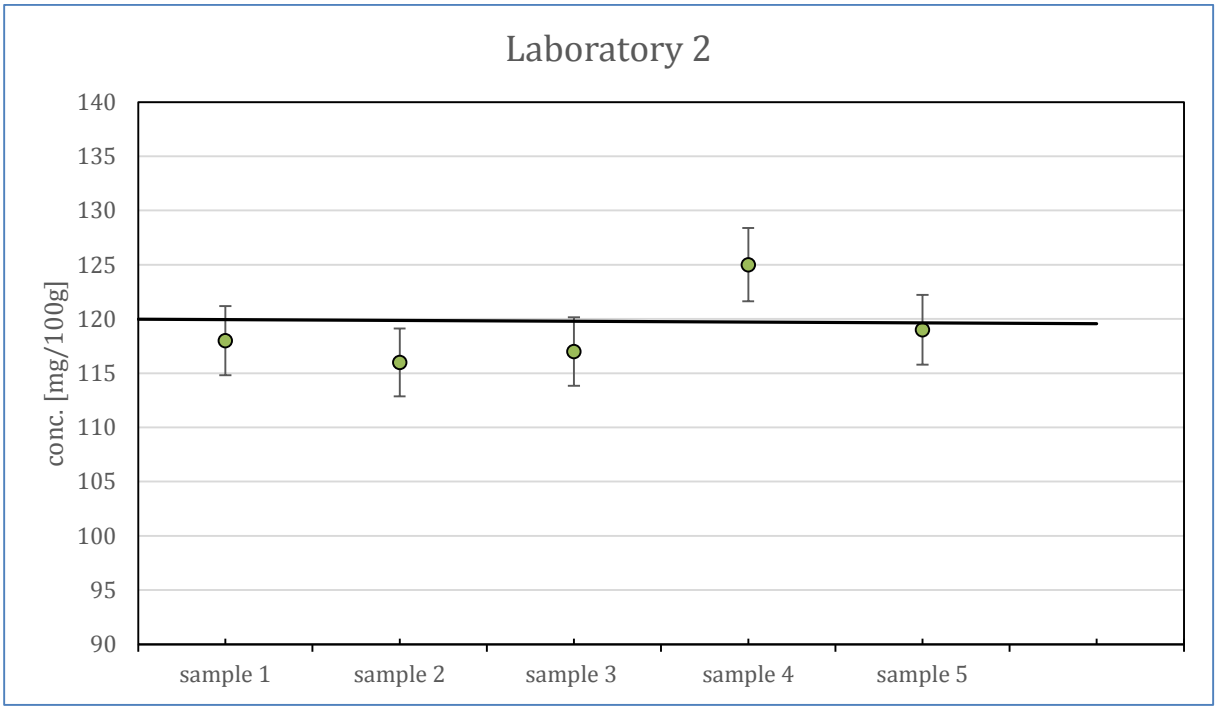

Figure 7: Laboratory 2 results

As a consequence, the result of Laboratory 1 would allow accepting the lot because no item of the samples contained more than 120 mg/100g with a sufficient probability. The result of Laboratory 2 would suggest refusing the lot because one of the samples contained more sodium than the maximum level with a probability higher than the required 95%.

For **inspection by variables (packages)** (12) measurement uncertainty might have an effect on the sample size. In the case that the measurement uncertainty σ_m is significant (higher than one tenth of

the sampling standard deviation s or process standard deviation σ), the sample size n should be increased by the factor $(1+\gamma^2)$ where $\gamma = \sigma_m/s$ or $\gamma = \sigma_m/\sigma$ respectively (12) (Annex P).

Example:
A lot of 500 items of pre-packaged Mineral Water is to be assessed for sodium content. For a given and agreed Acceptance Quality Limit AQL of 2.5% (maximum concentration 200 mg/L), usually (measurement uncertainty not significant) 30 samples should be taken for assessment (ISO 3951-2, Annex A, Table A1 and Annex B, Table B1). The production is well under control and the control charts give a process standard deviation σ of 2 mg/L. The measurement uncertainty standard deviation of the assessment laboratory σ_m is 1 mg/L, being therefore significant. With $\gamma = \sigma_m/\sigma = 0.5$ and $(1+\gamma^2) = 1.25$ the samples size is to be increased to 38. However, the decision parameter of the acceptability constant *k* is not to be altered.

For **inspection by variables (bulk)** (13), measurement uncertainty has an effect on the number of test samples per composite sample, n_T, and the number of measurements per test sample, n_M, when it is dominant. That is given when both the sampling increment standard deviation, σ_I, and the standard deviation between test samples, σ_P, are far less (one tenth or less) than the measurement standard deviation, σ_M, which is known and stable, (13) (Annex B). The number of sample increments n_I to form the composite samples (usually two) does not change in any case.

Example:

A lot of wheat bulk material is to be assessed for cadmium content (maximum concentration e.g. 0.1 mg/kg). Since cadmium is an ubiquitous contaminant, the concentration in the lot is very homogeneously distributed, giving very low standard deviations σ_I and σ_P, each estimated to 0.002 mg/kg. On the contrary, the analysed concentrations are very low, giving rise to a higher measurement uncertainty with a standard deviation σ_M of 0.02 mg/kg, being therefore dominant. The combined over all standard deviation $\sigma_O \approx 0.02$ mg/kg (square root of the sum of σ_M^2, σ^2 and σ^2) is divided by the discrimination interval D (difference between agreed risk-based acceptance and rejection level, here assumed to be 0.01 mg/kg) giving the relative standard deviation $d_O = \sigma_O / D \approx 2$. That parameter d_O is used (ISO 10725, Annex B, Table B1) to estimate the number of test samples per composite sample, $n_T = 6$, and the number of measurements per test sample, $n_M = 3$ (i.e. the analytical effort $n_T \cdot n_M = 18$). In the case that the measurement uncertainty would not have been dominant, the numbers would have been $n_T = 1$ and $n_M = 2$ (i.e. the analytical effort $n_T \cdot n_M = 2$).

Obviously, an additional laboratory workload lowering the measurement uncertainty would significantly reduce the analytical effort.

Final Remarks

By reflecting the word “uncertainty”, it does not necessarily generate a feeling of confidence.

However, in technical or scientific areas, this uncertainty specifies a characteristic value, which will be added to the result of a measurement. If the uncertainty value has been calculated according to a standardised procedure, the value reflects a particular power to have confidence in the measurement result.

By metrological definition, measurement uncertainty is a quantitative measure for quality regarding the actual result of the measurement. By this, it reflects how well the result is in line with the value of the measurand and enables the operator to estimate the reliability of the measurement and its established results. This is important while comparing more than just one measurement result or aligning to a reference material or a decision limit. Confidence in measurement results are helpful trading internationally and contributes to avoid trading constrains.

Literature

1. ***L. Kelvin, “Electrical Units of Measurement,” a Lecture Given on 3 May 1883, Published in the Book “Popular Lectures and Addresses, Volume 1,” 1891.***

2. International vocabulary of metrology — Basic and general concepts and associated terms (VIM), JCGM 200:2012.

3. Commission, European, [Hrsg.]. Poul-Erik Hansen, Gert Roebben, Introductory Guide to Nanometrology. 2010. ISBN: 978-0-9566809-1-4.
4. Estimation of measurement uncertainty in chemical analysis - Sisu@UT, available at https://sisu.ut.ee/measurement/uncertainty.
5. EURACHEM/CITAC Guide on the Use of uncertainty information in compliance assessment EURACHEM, 2007. This is available as a free download from http://www.eurachem.org/.
6. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories.
7. Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM), JCGM 100:2008 .
8. S L R Ellison and A Williams (Eds). Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, Third edition, (2012).
9. Eurolab technical Report No. 1/2007. Measurement Uncertainty Revisited: Alternative Approaches to Uncertainty Evaluation. Available as a free download from www.eurolab.org.
10. Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. B. Magnusson, T. Näykki, H. Hovind, M. Krysell. Nordtest technical report 537, ed. 3. Nordtest, 2011.
11. ISO 2859-1:1999 + Cor. 1:2001 + Amd.1:2011 Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.
12. ISO 3951-2:2010 Sampling procedures for inspection by variables -- Part 2: General specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection of independent quality characteristics, revised 2013 as ISO 3951.
13. ISO 10725:2000: Acceptance sampling plans and procedures for the inspection of bulk materials.
14. Introductory course on measurement uncertainty available on https://sisu.ut.ee/measurement/uncertainty.
15. Jülicher et al., Analyst, 1998, 123, S. 173-179.
16. Jülicher et al., Analyst, 1999, 124, 537-545.
17. ISO STANDARD 13528:2015 Statistical methods for use in proficiency testing by interlaboratory comparisons.
18. 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 Eurachem (2007).
19. Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method, JCGM 101:2008.
20. ISO 5725-1:1994 (trueness and precision) of measurement methods and results -- P: General principles and definitionsart 1. ***ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions.***
21. ISO 5725-2:1994 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.
22. ISO 5725-3:1994 Accuracy (trueness and precision) of measurement methods and results - Part 3: Intermediate measures of the precision of a standard measurement method.
23. ISO 5725-6:1994 Accuracy (trueness and precision) of measurement methods and results - Part 6: Use in practice of accuracy values.

24. ISO 21748:2010 (revised by ISO 21748:2017), "Guidance for the Use of Repeatability, Reproducibility and Trueness estimates in Measurement Uncertainty Estimation", ISO, Geneva (2010).
25. IUPAC/ISO/AOAC International Harmonized Guideline (7) (Protocol for the Design, Conduct and Interpretation of Method Performance Studies. Pure &Appl. Chem., Vol. 67, No. 2, pp. 331-343, 1995.),.
26. Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis, AOAC Official Methods Program, J. AOAC Int. 78(5), 143A–160A(1995) .
27. ***ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions.***
28. ISO STANDARD 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons.
29. Eurachem/Citac Guide „Quantifying uncertainty in analytical measurement".
30. ISO STANDARD 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons.
31. Analytical Methods Committee of the Royal Society of Chemistry "Uncertainty of Measurement - Implications of its use in Analytical Science", Analyst, 1995, 120 (9), 2303-2308.
32. NIST Technical note 1297 (1994 Edition): "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results".
33. NMKL Procedure No. 5, 2nd edition (2003): "Estimation and Expression of Measurement Uncertainty in Chemical Analysis".
34. UKAS (United Kingdom Accreditation Service) 2000 The Expression of Uncertainty in Testing Edition 1, UKAS Publication ref: LAB 12.
35. "Precision of Test Methods", Geneva, 1994, ISO 5725, Previous editions were issued in 1981 and 1986. (not adopted by Codex).
36. "Protocol for the Design, Conduct and Interpretation of Method Performance Studies", ed. W. Horwitz, Pure Appl. Chem., 1995, 67, 33 1-343 (adopted by Codex).
37. European Commission Decision 2002/657/EC implementing directive 96/23/EC Concerning the Performance of Analytical Methods and the Interpretation of Results, Off J Eur Comm, L22 1 (2002) 8-36.
38. Validation of Chemical Analytical Methods. NMKL Procedure No 4, 4th Version, 2010.
39. Eurolab Technical Report No 1/2007 Available at http://www.eurolab.org/.
40. Menditto, et al Accred. Qual. Assur. 2007, 12, 45.

APPENDIX III

Information Document Example Procedures for Estimating Measurement Uncertainty

Introduction

Every measurement comes with a particular imprecision. The quality of a measurement result is greatly improved if it comes with the estimation of a measurement uncertainty.

Measurement uncertainties is subject to the operator, the instrument used, the environment and any other sources, which may influences the measurement by a certain degree. When the uncertainty in a measurement is evaluated and stated, confidence in data obtained.

Such uncertainties can be evaluated and calculated upon by analysis of the measurement process. In practise, the total measurement uncertainty is usually calculated by combining several uncertainty contributions. There are established rules on how to calculate measurement uncertainty and guidelines are published to aid the undertaking.

The aim of this information document is to give some examples on the procedures for estimating measurement uncertainty and provide the reader with some references on the general topic.

Measurement procedures

In analytical chemistry, each measurement procedure can be subdivided into subsampling, subsample preparation, sample preparation, clean-up, calibration, quantification of the analyte and finally data analysis with evaluation of the measurement result (8). Figure 8 visualises the single steps:

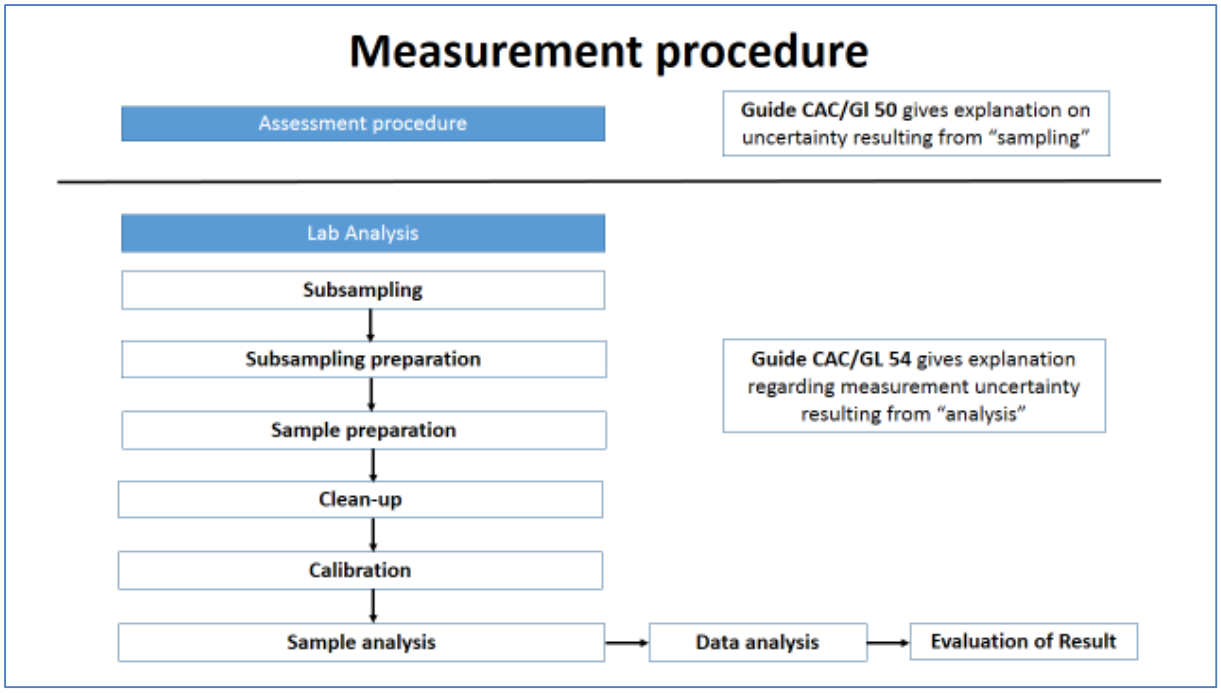

Figure 8: General steps for measurement procedure

Subsampling and subsampling preparation: A mostly prescribed procedure of taking parts from each of the samples, which had been taken from the lot according to the particular sampling plan.

Sample preparation: Most of the subsamples to be measured require treatment before they are going to be analyzed. Freezing, homogenization, dilution and extraction are only some of procedures mentioned. In many cases, analytes have to be converted into measurable compounds (e.g. a colourless sample is converted by addition of defined substances into a colored sample, which can be detected by UV-VIS spectroscopy). Due to possible decomposition or incomplete reaction, that "indirect" method might lead to loss of material or information. Additionally, contamination might also take place at any stage of the procedure and has to be avoided and controlled by analysing blank samples in parallel.

Calibration of analytical systems: In most cases, analyte-response curves need to be established from which the amount of analyte in question can be determined. Values has to be assigned by known substances, e.g. certified reference materials. As a consequence, the purity of the reference material and the any further prepared solution are influencing the measurement uncertainty.

Sample measurement: When the sample is finally measured, interferences of the analyte-defined substance complex with other components (e.g. reagents, matrix) can happen. The experience of the operators can have an impact on the measurement result. Instrument settings as well as the limited stability of the measurement device might cause various results and should be taken into account.

Data analysis: Processing algorithms (mathematical models, which are used to evaluate the results e.g. regression functions used for calibration) might differ from instrument to instrument.

Evaluation of the result: Analysis of the data, statistically evaluated can vary depending on the model used. Rounding as well as averaging of the final result should not expand the unavoidable measurement uncertainty.

Possible Uncertainty Sources

Subsampling:	Representative selection of parts of a lot sample
Storage/transportation:	special storage or shipping conditions with changing environmental condition
Instrumental effects:	detection limits, temperature, gas-pressure controller, gas flow regulator, auto-sampler with possible carry-over effects, time effects (measurement at various time points)
Purity/homogeneity:	partly-inhomogeneous samples, impure substances e.g. reagents, solutions or other used products
Measurement conditions:	Measurement of volumes: volumetric glassware effects for preparing solutions, various masses from different weightings; temperature effects; environmental changes e.g. humidity
Computational effects:	inaccurate calibration models, fitting procedures, rounding procedures
Blank correction:	Like sample, correction for Blank is necessary
Random effects:	By chance for all determination, should be included as a matter of course
Systematic effects:	Operator (experienced, unexperienced)

Table 1: Possible uncertainty sources

It might be noticed that not all possible sources of uncertainty will equally account for the uncertainty. In practice, it is likely that only a small number of all possible sources contribute significantly to the uncertainty. Unless there is a large number of contributions, components that are less than one third of the largest need not be evaluated in detail. (EURACHEM step3, 7.2.2.)

Procedures for estimating Measurement Uncertainty

Estimation of measurement uncertainty can be conducted by two main strategies:

The “bottom-up” approach, which determines the measurement uncertainty component by component. Every single source of error/uncertainty is separately estimated.

The “top-down” approach where the measurement uncertainty is estimated via error/uncertainty sources based on method performance data, e.g. validation studies, PTs etc.

These approaches refer to different situations:

Modelling (Classical ISO GUM)

- Uncertainty of an individual result of a measurement can be obtained, linked to a particular sample

Single-lab validation

- Typical uncertainty of results obtained using a defined procedure in the laboratory

Interlaboratory validation

- Uncertainty of results obtained using the same procedure in different laboratories

The modelling approach calculates the uncertainty for the individual result, on one concrete sample, for one situation. The single-lab validation approach is not linked to a particular sample; it is linked to a procedure. The interlaboratory approaches obtaining uncertainty results from the same procedure used in different laboratories. This type of approach gives some general uncertainty which can be expected when used in different laboratories.

Whichever of these approaches are going to be used, most of the information to calculate the measurement uncertainty is already available from previous studies done in order to validate existing or new methods, QA/QC data or studies, which has been carried out to test laboratory performances (14). Usually there is a lot of data available from proficiency testing data (PT), control chart data, calibration data from instruments, in-house validation data etc. The question remains, how to make the maximum use of these collected data to estimate the measurement uncertainty?

All starts with the definition of the measurand. Specifying a measurand is per se not an easy target. A clear definition of a.) an analyzed item or b.) a studied parameter is needed. For example, if the mass fraction of a residue (pesticide) is to be measured in a batch of two kilo gramms or in a piece of a single apple. It also makes a difference if a total amount of a heavy metal ion should be determined or the amount of its water-soluble salt. If the measurand has been clearly defined, the uncertainty operation can be distinguished between a “single laboratory” approach and “interlaboratory” approach.

Single laboratory approaches can be precede by model base or non-model base whereas in the first case component by component is evaluated. This type of practice is laid out in the ISO GUM and is considered as the standard approach for measurement uncertainty. Here the procedure is carefully analyzed, uncertainty sources are looked at component-by-component and then separately quantified. On the other hand, there is the non-model single laboratory described by Nordtest TR537 as well as a single laboratory approach, which includes an orthogonal design of experiments. This type of configuration is based on a statistical model (15) (16).

Interlaboratory approaches differ from single laboratory approaches. The interlaboratory approaches examine accumulated data from many laboratories, where each single laboratory does not contribute as much value to the final result. The situation or data from a single laboratory is not looked at and therefore this particular procedure is not highly recommended to establish measurement uncertainty. However, if the uncertainty has to be established for the first time or to preliminary find out more or less what the uncertainty will be this type of procedure can be used. For this purpose, ISO 13528:2015 on “Statistical methods for use in proficiency testing by interlaboratory comparison” describes procedures for robust data analysis (17).

In order to consider as many analytical situations as possible, the procedures are developed for different types of analytical methods (standard or in-house methods). Multi-factor experimental designs, analyzed by ANOVA, and Propagation of distributions using a Monte Carlo method are not included in this document but reference to literature is provided (18) (19) (15) (16).

This information document does not provide exemplary numerical calculations. It is assumed, that the concerned laboratories do have much experience on application of formulas.

Example Procedures for Estimating Measurement Uncertainty

The following procedures for estimation of measurement uncertainty should be regarded as practical examples, which are applicable in many day-to-day situations. They are not prescriptive. To achieve acceptance by both trading partners, the concepts are strictly based on internationally recommended guidelines and standards (JCGM 100:2008: Evaluation of measurement data - Guide to the expression of uncertainty in measurement (GUM) (7), the EURACHEM / CITAC Guide CG 4: Quantifying Uncertainty in Analytical Measurement (8) and ISO Protocols (20) (21) (22) (23) (24).

The relationship between measurement uncertainty of test samples, lot conformity assessment and sampling plans is explained in the Guidelines on measurement uncertainty (CAC/GL 54-2004).

The development of examples cannot be exhaustive and in special situations, other rational procedures might be applied by agreement. Nevertheless, they do not apply when legal specifications or other

internationally accepted guidelines define special rules for the estimation of the measurement uncertainty (e.g. the empirical Thomson-Horwitz equation). In particular, for pesticide residues, the procedures do not infringe on provisions in the Guidelines on estimation of uncertainty of results (CAC/GL 59-2006).

Measurement uncertainty, which is a parameter of the test result, is based on precision data of the method, taking into account the steps of analysis that may include sub-sampling, sample processing and instrumental analysis. The uncertainty components are combined according to the error propagation rules. Basically, N uncertainty standard deviations $s_{1...N}$ (or relative standard deviations i.e. coefficients of variation $cv_{1...N}$) of the statistical analysis can be combined to the total standard uncertainty u (or relative total standard uncertainty u_{rel}) (GUM 5.1.2, 5.1.5, 5.1.6) :

$$u = \sqrt{s_1^2 + s_2^2 + \ldots . + s_N^2} \qquad \text{or} \qquad u_{rel} = \sqrt{cv_1^2 + cv_2^2 + \ldots . + cv_N^2} \quad *)$$

*) The formulas refer to measurands given by the sum and/or the difference of parameters (left) or given by the product and/or the quotient of parameters (right). Since in practice, most of the analytical measurands are given by formulas with products and/or quotients of parameters, in the following text the second formula will be used. For simplicity, the parameters are regarded as non-correlated.

This provides the practical advantage that particular precision data from Single-Laboratory method validation or from inter-laboratory method validation (after proving fitness for purpose of the particular test laboratory by verification of that precision data) can be used in combination.

The following procedures are ordered by the particular type of the analytical method:

Type I:

- Defining Methods with additional consideration of subsample inhomogeneity and sample preparation variability

Type II:

- Rational Methods (Reference Methods)

Type III:

- Single Laboratory validated Methods (Alternative Approved Methods)
- Combination of repeatability precision of all single steps of analysis
- Precision estimated by series of analysis
- ISO 5752-2 and 5752-3 Approach
- Duplicate Approach

Type IV:

- Tentative methods: Ad-hoc Methods

Type I:

For Standard methods, the uncertainty is established utilising appropriate validation including precision data. Generally, these data are based on extensive inter-laboratory method validation, mostly performed according to the IUPAC/ISO/AOAC International Harmonized Guideline (25), ISO 5725-6 (currently under revision) or the AOAC International Guidelines for Collaborative Study Procedures to Validate Characteristics of a Method of Analysis (26).

A basic assumption underlying ISO 5725-1 (currently under revision) is that, for a standard measurement method, repeatability will be not be the same for all laboratories applying the standard procedure. However, the repeatability will be at least approximately the same, so that it is permissible to establish one common average repeatability standard deviation $s_{r,}$ which will be applicable to any laboratory, even if this is not 100% corresponding to the repeatability of the individual laboratory. Any laboratory should be carrying out a series of measurements under repeatability conditions and verify that the average repeatability standard deviation is applicable under given conditions (ISO 5725-6 (currently under revision)).

The reproducibility standard deviation s_R of the standard method is obtained by combining s_r with the between-laboratory standard deviation sL (ISO 5725-2 (currently under revision)).

Defining Methods with additional consideration of subsample inhomogeneity and sample preparation variability

Defining methods achieve comparability between laboratories measuring the same material with no intent to obtain an absolute measure of the true amount of analyte present. Corrections for method bias or matrix effect are ignored by convention.

If collaborative trial data are available, at least the repeatability should be evaluated in the particular laboratory and proven to be comparable to that s_r predicted by the collaborative trial and documented in the method i.e. the repeatability standard deviation should be less or equal s_r (EURACHEM Example A6).

A priori, no bias contribution must be considered and it is therefore appropriate to use the relative reproducibility standard deviation (i.e. the coefficient of variation) CV_R values from the collaborative trial or method publication as relative standard uncertainty u_{rel} within the tested range of analyte levels (EURACHEM 7.6.3).

Collaborative trials provide homogenized mostly stabilised material and hence do not cover physical preparation steps (e.g. grinding, drying) of the material. The uncertainty contributions of that analytical part should be additionally taken into consideration (EURACHEM 7.6.1), provided that the contribution is significant (i.e. >1/3 CV_R (EURACHEM 7.2.2)).

In the case of significant laboratory sample inhomogeneity, the uncertainty contribution of subsampling should be considered. The significance might be assessed by using a homogeneity check like ISO 13528 (17), Annex B by comparing the relative between-subsamples standard deviation cv_s with the relative standard deviation for proficiency assessment CV_σ (σ is used for the estimation of the z-scores) of the standard method. The laboratory sample may be considered to be adequately homogeneous if, $cv_s \leq 0.3\ CV_\sigma$.

The between-subsamples standard deviation s_s might be estimated by the procedure given in ISO 13528, Annex B1 and using the formula given in Annex B3. That duplicate test gives information also on the uncertainty contribution of the physical preparation procedure:

Select a number g of the subsamples from the laboratory sample at random, where g ≥10.

- Prepare two test portions from each subsample using techniques appropriate to the test material to minimize between-test-portion differences.
- Taking the 2g test portions in a random order, obtain a measurement result on each, completing the whole series of measurements under repeatability conditions.
- Calculate the general average $\bar{\bar{x}}$

$$\bar{\bar{x}} = \frac{\sum_{t=1}^{g} \bar{x}_t}{g} \qquad \text{with} \qquad \bar{x}_t = \frac{x_{t,1} + x_{t,2}}{2}$$

- Calculate the standard deviation s_x of sample averages

$$s_x = \sqrt{\frac{\sum_{t=1}^{g} (\bar{x}_t - \bar{\bar{x}})^2}{g - 1}}$$

- Calculate the within-subsamples standard deviation s_w which is a measure of the physical preparation uncertainty

$$s_w = \sqrt{\frac{\sum_{t=1}^{g} w_t^2}{2g}} \qquad \text{with} \qquad w_t = \left| x_{t,1} - x_{t,2} \right|$$

- Calculate the between-subsamples standard deviation s_s with the factor ½ on s_w due to the mean of duplicate analyses being used

$$s_S = \sqrt{s_x^2 - \frac{s_W^2}{2}}$$

- and the relative standard deviation of sample inhomogeneity

$$cv_S = \frac{s_S}{\bar{x}}$$

In case that the sample inhomogeneity is significant ($cv_s > 0.3\ CV_\sigma$), the relative standard measurement uncertainty u_{rel} is given by the combination:

$$u_{rel} = \sqrt{cv_R^2 + cv_S^2}$$

Taking into account the uncertainty contribution of sample preparation (the standard deviation is divided by √2 to correct from a standard deviation for pairwise differences to the standard uncertainty for single values),

$$cv_P = \frac{1}{\sqrt{2}} \frac{s_w}{\bar{x}}$$

the relative standard measurement uncertainty u_{rel} is given by the combination:

$$u_{rel} = \sqrt{cv_R^2 + cv_S^2 + cv_P^2}$$

Note: In formulas for calculating the analytical result, the influence of subsampling differences due to inhomogeneity and preparation variability can be implemented as factors, which are dispersed around 1 (EURACHEM A4.3).

Type II:

Rational Methods (Reference Methods)

For rational standard methods, trueness is an issue, which should be considered in the estimation of measurement uncertainty. The current procedure applies to the situation where no bias is to be taken into account. But this assumption should be proven by appropriate recovery experiments.

For many rational standard methods, certified reference materials are supplied. As an alternative, samples can be spiked with a known level of the analyte (with preference of matrices, which do not contain the analyte), bearing in mind the different behaviour of the spiked substance and the native counterpart.

In a first step, from n recovery experiments on certified reference material or homogenized spiked material (e.g. homogenized samples are split and one portion spiked) with the reference concentration x_{ref} , the found concentrations of the analyte x_i , and the bias b_i , the average laboratory bias $\bar{b}$ is estimated

$$\bar{b} = \frac{1}{n}\sum_{i=1}^{n} b_i \qquad \text{with} \qquad b_i = x_i - x_{ref}$$

and compared with the standard uncertainty u at the reference concentration (by multiplying u_{rel} with the concentration of the analyte) combined with the certified uncertainty of the reference material or the experimental uncertainty of spiked material estimated by homogeneity tests u_{ref} (see 4.1.1). Laboratory bias can be neglected if

$$\left|\bar{b}\right| \leq 2\sqrt{\left(\frac{u^2}{n}\right) + u_{ref}^2}$$

Otherwise, the bias is significant (EURACHEM 7.16) and the analytical result might be corrected for the bias, making due allowance for the uncertainty of the correction. In that case, the standard deviation s_B of the average bias is given by

$$s_B = \frac{1}{\sqrt{n}}\sqrt{\frac{\sum_{i=1}^{n}\left(b_i - \bar{b}\right)^2}{n-1}}$$

In case that the matrix might have an impact on the bias, the recovery experiments should be applied on samples from different matrices and the uncertainty contribution of that particular matrix, which corresponds to the sample should be used.

Note: It should be avoided to take the effect of bias (this is not the uncertainty of bias) into account by enlarging the “uncertainty” assigned to the result instead of correcting for bias. Evaluating the uncertainty of a measurement result should not be confused with assigning a safety limit to some quantity (Guide to the expression of uncertainty in measurement (GUM), 6.3.1).

Type III:

Single-laboratory Validated Methods (Alternative Approved Methods)

Contrary to standard methods, for Single-laboratory validated methods no published standard precision data are available. Therefore, they are subjects of extensive validation procedures. Despite of ad-hoc situations, the validation provides precision data.

In case that the Single-laboratory validated method is a modification of a corresponding standard method, the estimation of precision should focus on the uncertainty contributions of that modification. The uncertainty contributions should be compared to the relative reproducibility standard deviation (i.e. coefficient of variation) CV_R values from the collaborative trial or standard method publication. If the uncertainty contribution of modifications is negligible, it is appropriate to use CV_R as relative standard uncertainty u_{rel} and to proceed according to Procedures 4.1.

There are two general approaches to estimate the precision:

- The combination of the repeatability precision of all single steps of analysis (e.g. weighing, drying, extracting, diluting and analytical measurement) with the involved calibrations and other uncertainty sources (e.g. purity of reference standards, experience of test personnel)
- Precision estimated by series of analysis as far as possible over an extended time period allowing natural variation of all impact factors.

In practice, a combination of these types is usually necessary and convenient. Therefore, a variance component model offers the possibility of covering various components of the overall uncertainty within one validation experiment, including a randomised sampling scheme (16).

Combination of the repeatability precision of all single steps of analysis

The uncertainty components associated with N potential sources of uncertainty are identified, quantified as standard deviations u_i, multiplied with sensitivity coefficients c_i, and combined (GUM 5.1.3):

$$u = \sqrt{\sum_{i=1}^{N} c_i \cdot u_i^2}$$

Note: In the case that the different components are not statistically independent, corresponding correlation factors are to be introduced.

The sources are for example:

- Standard substances (certified uncertainty/purity)
- Physical/chemical variability (extraction, derivatisation, stoichiometry)
- Application of measuring devices for preparation of the test samples (balances, pipettes, thermometers etc.)
- Application of analytical instruments (stability, calibration, contamination etc.)
- Different experience of staff

The procedure begins with the critical reflection of the formula of the measurand i.e. the relationship between the result and the input values. All parameters are to be checked for their uncertainty relevance.

Therefore, for example, the uncertainty of the sample preparation is separated into the uncertainties of the individual steps of weighing, homogenizing, drying, extracting, diluting etc., which are to be combined.

The uncertainty of weighing itself, for example, is estimated from the separate contributions of calibration and traceability (including certified uncertainty of the weights) and the uncertainty of the reading (analogue/digital-display).

Obviously, the subject of this type of estimation is too complex to be sufficiently described in the current paper. Therefore, for further information, reference is made to the JCGM 100:2008: Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM) and the EURACHEM / CITAC Guide CG 4: Quantifying Uncertainty in Analytical Measurement.

Precision estimated by series of analysis

According to ISO 5725-3, precision estimated in one laboratory is the so-called intermediate precision measure, which is usually smaller than the reproducibility standard deviation based on inter-laboratory method validation and hence more appropriate for the individual laboratory. That intermediate precision condition of measurement includes the same measurement procedure, same location, and replicate measurements on the same or similar objects over an extended period of time, but may include other conditions involving changes like new calibrations, calibrators, operators, and measuring systems.

Estimation of precision should take into account all parts of the analysis, which basically would be involved in case of participation on a corresponding inter-laboratory validation of a standard method. That comprises at least the extraction/derivatisation/digestion procedures, which could possibly lead to recovery variation. The complete measurement process also includes calibration and traceability.

A typical test sample containing an appropriate amount of analyte (e.g. homogenized and dried or processed to assure stability of the matrix and analyte(s)) might be analyzed several times over a period of time, using different analysts and equipment where possible (e.g. the results of measurements on quality control samples) thus verifying Single-Laboratory reproducibility conditions (EURACHEM 7.7.2) or intermediate precision conditions.

The relative intermediate standard deviation cv_{int} estimated by use of the following procedures, like corresponding collaborative trials, does not cover effects of sample preparation and subsample inhomogeneity. In order to take into account these uncertainty components, they should be combined with cv_{int} as described in Procedures 2.1.

For the identification and uncertainty estimation of bias, the approaches described in the Procedure 2.1.2 have to be applied.

In case that the uncertainty might depend on analyte levels, the precision experiments should be carried out at different levels in any case including the level, which is relevant for compliance assessment. The significance of influence might be checked by the F-test or the Cochran test for homogeneity of the variances from different experiments on different levels of the analyte.

Finally, the uncertainty of the calibration standards (which obviously might be much higher than the certified uncertainty of reference material) or of the reference materials (negligible in most cases) should be considered.

ISO 5725-2 and ISO 5725-3 Approach

An appropriate norm-consistent approach might be the as-far-as-possible-application of the procedure given in ISO 5725-2 and ISO 5725-3. The reproducibility standard deviation s_R of an inter-laboratory method validation is obtained by combining the mean repeatability standard deviation s_r of all laboratories with the between-laboratory standard deviation s_L.

A typical test sample (homogenized and dried) is analyzed over a period of time on n different days by different analysts (with a new extraction/digestion, recalibration). Each of the days, a number of k replicates of the particular extract/digest are measured with the results $x_{j=1...k}$ under repeatability conditions (measurement within a short time, the same instrument and calibration used by the same operator) and the following parameters are calculated:

- Each day i : From the k replicate results $x_{j=1...k}$ the mean value $\bar{x}_i$ and the repeatability standard deviation $s_{r\,i}$ are estimated:

$$\bar{x}_i = \frac{1}{k}\sum_{j=1}^{k} x_j$$

$$s_{r\,i} = \sqrt{\frac{\sum_{j=1}^{k}\left(x_j - \bar{x}_i\right)^2}{k-1}}$$

- From the repeatability standard deviations of the different days $s_{r\,i=1...n}$, the mean repeatability standard deviation $s_{r\,mean}$ is calculated:

$$s_{r\,mean} = \sqrt{\frac{\sum_{i=1}^{n} s_{r\,i}^2}{n}}$$

- The "between-days" standard deviation s_d of the mean values $\bar{x}_{i=1...n}$ of the different days is calculated:

$$s_d = \sqrt{\frac{\sum_{i=1}^{n}(\bar{x}_i - \bar{\bar{x}})^2}{n-1}}$$

with the total mean value $\bar{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n}\bar{x}_i$

- According to ISO 5725-3, the intermediate standard deviation is given by :

$$s_{int} = \sqrt{s_{r\,mean}^2 + s_d^2}$$

Finally, the relative intermediate standard deviation is given by:

$$cv_{int} = \frac{s_{int}}{\bar{\bar{x}}}$$

Duplicate Approach

As an alternative to the above-mentioned ISO 5725-2 and ISO 5725-3 approach, the overall run-to-run variation can be performed with a number n of duplicate tests (homogenized samples each divided into two test samples, each of the test samples subjected to complete extraction/digestion and determination procedure including recalibration)(EURACHEM 7.7.2 and A4.4).

For each duplicate test i, the relative differences δi rel and the standard deviation of the relative differences sδrel are calculated:

$$\delta_{i\,rel} = \frac{\delta_i}{\bar{x}_i}$$

with $\delta_i = x_{i,1} - x_{i,2}$ and $\bar{x}_i = \frac{x_{i,1} + x_{i,2}}{2}$

$$s_{\delta_{rel}} = \sqrt{\frac{\sum_{i=1}^{n}\left(\delta_{i\,rel} - \bar{\delta}_{rel}\right)^2}{n-1}}$$

with $\bar{\delta}_{rel} = \frac{1}{n}\sum_{i=1}^{n}\delta_{i\,rel}$

- Finally, this standard deviation is divided by √2 to correct from a standard deviation for pairwise differences to the standard uncertainty for single values giving the relative intermediate standard uncertainty:

$$cv_{int} = \frac{s_{\delta_{rel}}}{\sqrt{2}}$$

Type IV:

Ad-hoc Methods (Tentative Methods)

In most cases, ad-hoc methods are based on standard or well-established Single laboratory validated methods. They are expanded substantially (e.g. to other analytes or matrices) and will not generally require complete revalidation, but the procedure, which was described in the first paragraph of ISO 5725, Procedures 4.2 is highly recommended. Further information on the evaluation of the measurement uncertainty for ad-hoc methods are given in the EURACHEM Guide (EURACHEM 7.10). In order to get an acceptable statistical power, as many replicates as practical of the test (including all relevant parts of method) should be performed. The comparison of the resulting relative standard deviation with the relative standard uncertainty of the basic method gives information about the precision equivalence of the ad-hoc method. Where appropriate, the uncertainty of the basic method should be reported.

Literature

1. ***L. Kelvin, "Electrical Units of Measurement," a Lecture Given on 3 May 1883, Published in the Book "Popular Lectures and Addresses, Volume 1," 1891.***

2. International vocabulary of metrology — Basic and general concepts and associated terms (VIM), JCGM 200:2012.

3. Commission, European, [ed.]. Poul-Erik Hansen, Gert Roebben, Introductory Guide to Nanometrology. 2010. ISBN: 978-0-9566809-1-4.

4. Estimation of measurement uncertainty in chemical analysis - Sisu@UT, available at https://sisu.ut.ee/measurement/uncertainty.

5. EURACHEM/CITAC Guide on the Use of uncertainty information in compliance assessment EURACHEM, 2007. This is available as a free download from http://www.eurachem.org/.

6. ISO/IEC 17025:2005 General requirements for the competence of testing and calibration laboratories.

7. Evaluation of measurement data — Guide to the expression of uncertainty in measurement (GUM), JCGM 100:2008 .

8. S L R Ellison and A Williams (Eds). Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, Third edition, (2012).

9. Eurolab technical Report No. 1/2007. Measurement Uncertainty Revisited: Alternative Approaches to Uncertainty Evaluation. Available as a free download from www.eurolab.org.

10. Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories. B. Magnusson, T. Näykki, H. Hovind, M. Krysell. Nordtest technical report 537, ed. 3. Nordtest, 2011.

11. ISO 2859-1:1999 + Cor. 1:2001 + Amd.1:2011 Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.

12. ISO 3951-2:2010 Sampling procedures for inspection by variables -- Part 2: General specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection of independent quality characteristics, revised 2013 as ISO 3951.
13. ISO 10725:2000: Acceptance sampling plans and procedures for the inspection of bulk materials.
14. Introductory course on measurement uncertainty available on https://sisu.ut.ee/measurement/uncertainty.
15. Jülicher et al., Analyst, 1998, 123, S. 173-179.
16. Jülicher et al., Analyst, 1999, 124, 537-545.
17. ISO STANDARD 13528:2015 Statistical methods for use in proficiency testing by interlaboratory comparisons.
18. 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 Eurachem (2007).
19. Evaluation of measurement data — Supplement 1 to the "Guide to the expression of uncertainty in measurement" — Propagation of distributions using a Monte Carlo method, JCGM 101:2008.
20. ISO 5725-1:1994 (trueness and precision) of measurement methods and results -- P: General principles and definitionsart 1. ***ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions.***
21. ISO 5725-2:1994 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.
22. ISO 5725-3:1994 Accuracy (trueness and precision) of measurement methods and results - Part 3: Intermediate measures of the precision of a standard measurement method.
23. ISO 5725-6:1994 Accuracy (trueness and precision) of measurement methods and results - Part 6: Use in practice of accuracy values.
24. ISO 21748:2010 (revised by ISO 21748:2017), "Guidance for the Use of Repeatability, Reproducibility and Trueness estimates in Measurement Uncertainty Estimation", ISO, Geneva (2010).
25. IUPAC/ISO/AOAC International Harmonized Guideline (7) (Protocol for the Design, Conduct and Interpretation of Method Performance Studies. Pure &Appl. Chem., Vol. 67, No. 2, pp. 331-343, 1995.),.
26. Guidelines for Collaborative Study Procedures To Validate Characteristics of a Method of Analysis, AOAC Official Methods Program, J. AOAC Int. 78(5), 143A–160A(1995) .
27. ***ISO 5725-1:1994 Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions.***
28. ISO STANDARD 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons.
29. Eurachem/Citac Guide „Quantifying uncertainty in analytical measurement".
30. ISO STANDARD 13528:2005 Statistical methods for use in proficiency testing by interlaboratory comparisons.
31. Analytical Methods Committee of the Royal Society of Chemistry "Uncertainty of Measurement - Implications of its use in Analytical Science", Analyst, 1995, 120 (9), 2303-2308.
32. NIST Technical note 1297 (1994 Edition): "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results".

33. NMKL Procedure No. 5, 2nd edition (2003): "Estimation and Expression of Measurement Uncertainty in Chemical Analysis".
34. UKAS (United Kingdom Accreditation Service) 2000 The Expression of Uncertainty in Testing Edition 1, UKAS Publication ref: LAB 12.
35. "Precision of Test Methods", Geneva, 1994, ISO 5725, Previous editions were issued in 1981 and 1986. (not adopted by Codex).
36. "Protocol for the Design, Conduct and Interpretation of Method Performance Studies", ed. W. Horwitz, Pure Appl. Chem., 1995, 67, 33 1-343 (adopted by Codex).
37. European Commission Decision 2002/657/EC implementing directive 96/23/EC Concerning the Performance of Analytical Methods and the Interpretation of Results, Off J Eur Comm, L22 1 (2002) 8-36.
38. Validation of Chemical Analytical Methods. NMKL Procedure No 4, 4th Version, 2010.
39. Eurolab Technical Report No 1/2007 Available at http://www.eurolab.org/.
40. Menditto, et al Accred. Qual. Assur. 2007, 12, 45.

APPENDIX IV

LIST OF PARTICIPANTS

Argentina Gabriela Catalani Agroindustry Ministry	Australia Karina Budd Department of Agriculture and Water Resources Kate Slater Department of Agriculture and Water Resources Richard Coghlan National Measurement Institute
Brazil Carolina Araújo Vieira Brazilian Health Regulatory Agency Ligia Lindner Schreiner	Canada Barbara Lee Health Canada Thea Rawn Health Canada
Bulgaria Ivelin Rizov Ministry of Agriculture, Food and Forestry	Chile Catherine Caceres Insituto Salud Publica de Chile
Colombia Myriam Rivera Rico Invima Colombia Pais Miembro Codex	China Stephen Chung Wai-Cheung Centre for Food Safety
Dominican Republic Fátima del Rosario Cabrera Ministerio de Salud Pública y Asistencia Social	
European Union Franz Ulberth	Ecuador Víctor Hugo Almeida Arteaga Ministerio de Salud Pública del Ecuador
Egypt Mariam Barsoum Onsy Egyptian Organization for Standardization & Qyuali	France Bertrand Lombard ANSES Concordet Ecole Nationale Vétérinaire de Toulouse
Germany Katrin Franks Chair of EWG BVL	India Anoop A Krishnan Export Inspection Agency-Kochi Laboratory Krishna Kumar Sharma Indian Agricultural Research Institute Sunil Bakshi
Iran Samaneh Eghtedari ISIRI	Ireland Ita Kinahan The State Laboratory
Japan Hidetaka Kobayashi Ministry of Agriculture, Forestry and Fisheries Takahiro Watanabe National Institute of Health Sciences Yukiko Yamada Ministry of Agriculture, Forestry and Fisheries	Kazakhstan Zhanar Tolysbayeva The Ministry of Healthcare
Malaysia Nuurul Hidayah Sharipan Ministry of Health	México Emmanuel Hernández Galván

	Tania Daniela Fosado Soriano Secretaría de Economía
Netherlands Henk van der Schee NVWA Yannick Weesepoel RIKILT - Wageningen University and Research	Norway Stig Valdersnes Institute of Marine Research
Poland Magdalena Kowalska	Republic of Korea Chaehyung Kim Ministry of Food and Drug Safety
South Africa Ephraim Moruke Department of Agriculture, Forestry and Fisheries	Switzerland Erik Konings Nestec Ltd. Gerard Gremaud Swiss Federal Food Safety and Veterinary Office
Thailand Chanchai Jaengsawang Department of Medical Sciences Rungrassamee Mahakhaphong ACFS	Tunisia Samiha Toumi Unit of Pharmacy and Medicines (UPM), Ministry of Health
United Kingdom Chelvi Leonard Food Standards Agency	Uruguay Laura Flores LATU
USA Greg Noonan US FDA	International Fruit & Vegetable Juice Association John Collins
AACCI Terry C Nelsen	International Dairy Federation Aurelie Dubois Jaap Evers