

FOOD ANALYSIS: THE ROLE OF MEASUREMENT UNCERTAINTY INTERNATIONAL CONFERENCE 2ND IMEKOFOODS

RELIABILITY OF MEASUREMENT UNCERTAINTY ESTIMATES: CURRENT DEVELOPMENTS

Marina Patriarca¹, Barbara De Santis², Carlo Brera²

¹Istituto Superiore di Sanità, Methodologies and indicators for chemical safety in the food chain and human health Unit, Department of Veterinary Public Health and Food Safety, Rome, Italy, marina.patriarca@iss.it

²Istituto Superiore di Sanità, GMO and Mycotoxin Unit, Department of Veterinary Public Health and Food Safety, Rome, Italy, barbara.desantis@iss.it, carlo.brera@iss.it

Abstract – Since the publication of the “Guide to the estimate of measurement uncertainty”, in 1993, several efforts were put in place to implement the principles of uncertainty estimate in all fields of measurement and testing. In the area of chemical analysis, including food analysis, it was early recognised that a purely metrological approach was not fully applicable. Eurachem and other organizations played a major role in implementing alternative approaches based on collaborative studies and the wealth of data already available, even to single laboratories, from quality assurance practices. More recently, given the increasing use of measurement uncertainty in compliance statements, attention has been paid to the setting of target measurement uncertainty as a key requisite for the validation of analytical methods. Finally, recent guidance recommends to providers of proficiency testing schemes to alert their participants regarding the plausibility of their reported uncertainty estimates. These aspects will be explored with regards to food analysis.

Keywords: measurement uncertainty, validation studies, collaborative studies, target uncertainty, proficiency testing, food analysis

1. INTRODUCTION

Measurement uncertainty is increasingly being accepted as the key parameter to determine the level of confidence that can be placed on a measurement result and therefore it is increasingly used as part of compliance assessments. It is

therefore of the utmost importance that measurement uncertainty estimates are reliable.

In 1993 the “Guide to the expression of uncertainty in measurement (GUM)” [1] stated the fundamental principles on which such estimates should be based. A reissue of this document with minor corrections was made available in 2008 by the Joint Committee for Guides in Metrology [2] and is also endorsed as ISO/IEC Guide 98-3 [3]. These general rules for evaluating and expressing uncertainty in measurement can be followed at various levels of accuracy, in all sectors of measurement and testing, from everyday tasks to fundamental research. In addition, the Guide, acknowledging that a purely metrological approach could not be generally applied as such, also made provisions for reaching reliable measurement uncertainty statements based on previous knowledge and sometimes even expert judgement. Although such estimates are somewhat less rigorous than those based on detailed examination of the measurement process, they fulfil the task if they are reliable enough for the intended purpose.

Already in the year 2000, Eurachem exploited these concepts to provide detailed guidance on how to estimate measurement uncertainty in quantitative analytical measurements. This guidance was recently revised and expanded to cover new developments in the field [4].

An even more practical approach to the estimate of measurement uncertainty was developed by Magnusson et al. [5] for the benefit of busy routine analytical laboratories, using examples from environmental analysis. The so-called Nordtest approach, based on the results of carefully planned internal quality control and proficiency testing, was well received by both laboratories and accreditation

bodies and it proved useful in many fields of analytical measurements, including the food sector.

In many areas and particularly in food analysis, standard methods, whose performance in terms of repeatability and reproducibility, were tested in collaborative studies, had long been a stronghold in support of the comparability of measurement results obtained in different times and places. Given the effort spent in developing such methods, the International Organization for Standardization (ISO) undertook the task to assess how the existing performance data could be used to estimate the uncertainty of measurement results produced by those methods. The resulting guidance was published in 2010 [6].

However estimated, a measurement uncertainty is only useful if it is fit for the intended use. This issue was taken on board in the European Union (EU) legislation regarding contaminants in food and the concept of “maximum standard measurement uncertainty” was developed as a fitness for purpose approach for methods developed in-house by the laboratories. More recently, Eurachem issued new guidance addressing how to set target measurement uncertainty as a key requisite for the validation of analytical methods [7].

Finally, laboratories should not only estimate the measurement uncertainty of their results but also prove that their estimate is reliable. This is of the utmost importance in those areas, such as in food analysis, where measurement uncertainty plays an important role in compliance assessment. Participation in proficiency testing (PT) provides not only an assessment of the performance of the laboratory but also sufficient information for the laboratory to assess its own claim regarding measurement uncertainty. In addition the revised ISO 13528 [8] points out to providers of PT schemes the need to alert their participants regarding the plausibility of their uncertainty estimates.

These aspects will be explored in more detail.

2. • MEASUREMENT UNCERTAINTY ESTIMATES FROM METHOD PERFORMANCE DATA

It is widely recognised that, because of the difficulties associated with defining mathematical models for analytical measurements, a purely metrological approach, based on the individual assessment of the contribution of each input quantity and influence factor to the measurement uncertainty on the final result, may grossly

underestimate the uncertainty effectively associated with it. To overcome these problems, top-down approaches, relying mainly on data produced within the laboratory in its everyday practice and including information from method validation studies as well as results of the internal quality control and participation in PT, were proposed [4, 5] and are widely used. Guidance for estimating the uncertainty of standard methods from the performance data obtained in collaborative studies is given in ISO 21748 [6].

2.1. • Evaluation of measurement uncertainty from the method validation data

Laboratories developing their own test methods and/or adopting those developed by others (e.g. scientists or manufacturers) are requested to validate them, by determining key performance parameters under stated conditions, and comparing them successfully with pre-defined criteria for the fitness-for-purpose.

Typically, a validation plan will include an evaluation of the method precision, in terms of repeatability and intermediate precision, and an assessment of the bias (or, at least, lack of significant bias). Depending on the circumstances, determination of other parameters may be requested, such as limit of detection, limit of quantification, working range and robustness, but these do not normally play a significant role in the estimate of measurement uncertainty. Further guidance on the terminology used here and on how to plan and carry-out a validation study can be found in two other Eurachem Guides [9, 10].

The procedure for estimating measurement uncertainty as described in the GUM [1-3] starts from the definition of the measurand and the model equation to be used for its quantification. An initial analysis should be carried out to identify all sources of uncertainty potentially affecting the measurement results. The effects should be quantified, either individually or grouping effects that can be quantified together. To this aim, the best available estimate of overall precision and the best available estimate(s) of overall bias and its uncertainty cover the effects of most uncertainty sources on the measurement results. In many cases combining these two uncertainty components using the law of propagation of uncertainty provides a reasonable estimate of the standard measurement uncertainty of the results obtained when using the method. Care must be taken to include, in the

validation experiments, studies of the effects of the variation of all important sources of uncertainty identified and, when necessary, to consider also any important additional component not covered by them. A more detailed description of this process, including specific examples related to food analysis, is given in the Eurachem guide [4].

2.2. Evaluation of measurement uncertainty for a method in routine use

In some cases, methods may have been in use for some time before a formal evaluation of the measurement uncertainty associated with their results is undertaken. In addition, the laboratory may have performed an initial estimate of the measurement uncertainty, based on the data available, either with a top-down or a bottom-up approach, but may wish to re-evaluate it after some time, when more data on the method performance under routine use have become available. Typically, in such cases, the laboratory has collected data from the internal quality control, ideally using test samples, and may have participated to at least a few PT, covering the matrices within the scope of the methods or analysed additional certified reference materials (CRMs).

The Nordtest approach [5] provides a simple, although rough, approach to the estimate of measurement uncertainty, which major strength is the use of data representing the actual method performance on test samples during routine use and its drawback the need to rely on several data, collected over a sufficient period of time (typically at least one year).

The model behind this approach still relies on grouping most uncertainty contributions in two groups:

- a) those contributing to the random variability, estimated as the standard deviation observed for the quality control samples, and
- b) those arising from variable biases within the method, estimated from the deviations from assigned values observed analysing several CRMs or PT samples.

The general equation used to calculate the measurement uncertainty will be:

$$u_c = \sqrt{u(R_w)^2 + u(bias)^2}$$

where $u(R_w)$ represents the standard deviation observed from the analysis of quality control

samples and $u(bias)$ is obtained from the following equation:

where RMS_{bias} represents bias variability

$$RMS_{bias} = \sqrt{\frac{\sum (bias_i)^2}{n_{CRM}}}$$

and $\bar{u}(Cref)$, calculated as follows:

$$\bar{u}(Cref) = \sqrt{\frac{\sum u(Cref_i)^2}{n_{CRM}}}$$

takes into account the uncertainties of the assigned values of the CRMs or PT samples analysed.

2.3. Evaluation of measurement uncertainty for standard methods

ISO 21748 [6] provides guidance for the use of repeatability and reproducibility data obtained for standard methods in collaborative studies designed according to ISO 5725 [11]. Provided that certain conditions are met, it is recommended that laboratories report their expanded measurement uncertainty as twice the stated reproducibility standard deviation for that method, thus providing the basis for harmonised compliance judgements.

Laboratories should first obtain estimates of the repeatability, reproducibility and trueness of the method in use from published information about the method, then perform appropriate experiments to demonstrate that they can operate the method appropriately, obtaining repeatability and bias consistent with those stated for the method. These would normally consist in analysis of CRMs over the method working range. Experiment planning must ensure that sufficient degrees of freedom are achieved, to support meaningful conclusions from the statistical tests used to compare the laboratory performance with the stated method performance data.

In addition, it is the laboratory responsibility to identify any influences on the measurement that were not adequately covered in the collaborative study and undertake assessment to quantify any additional contribution to the uncertainty arising from these effects. Significant contributions, such as those arising from test sample inherent

$$u(bias) = \sqrt{RMS_{bias}^2 + \bar{u}(Cref)^2}$$

inhomogeneity, should be evaluated and combined with the reproducibility standard deviation, in the usual way according to the law of uncertainty propagation.

In horizontal methods for food analysis collaborative studies often cover several matrices and different concentrations. In order to calculate the measurement uncertainty it will be necessary to identify the model(s) describing the variation of the reproducibility over the range of concentrations and matrices and derive the appropriate value for the expected reproducibility standard deviation to be used in the calculation of the measurement uncertainty for an individual result.

3. FITNESS FOR PURPOSE AND PLAUSIBILITY OF THE ESTIMATES OF MEASUREMENT UNCERTAINTY

However estimated, a measurement uncertainty is only useful if it is fit for the intended use of the measurement result. Therefore, attempts are in place in legislation to set limits for “acceptable” uncertainty. More specifically, Eurachem examined different approaches to setting target measurement uncertainty as a key requisite for the validation of analytical methods depending on the type of answer requested from the analytical laboratory [7].

Ultimately, it is the laboratory responsibility to prove that their measurement uncertainty estimate is reliable, especially when, as in some areas of food analysis, measurement uncertainty plays an important role in compliance assessment. Participation in proficiency testing may help in this direction, providing participants not only with the means to assess the equivalence of their measurement results with those provided by other laboratories and even by other analytical techniques, but also the possibility to check the plausibility of the interval they claim to represent the possible values of the measurand.

3.1. Setting the target measurement uncertainty

A number of subjects (analysts, regulators and other end-users of analytical information) may have an interest in setting values for “target measurement uncertainty”, defined in the International vocabulary of metrology (VIM) [12] as “measurement uncertainty specified as an upper limit and decided on the basis of the intended use of measurement results”.

In some cases, when limits for the measurand value are set, guidance about the acceptable magnitude of the uncertainty may be described in legislation, technical specification or other guidelines referring to the assessment of compliance with those limits. In EU legislation regarding, for example, mycotoxins [13] and heavy metals [14] in food, the concept of “maximum standard measurement uncertainty” is applied as a fitness for purpose approach for methods developed in-house by the laboratories. The combined standard uncertainty of the measurement results produced by in-house methods should be less than the maximum standard measurement uncertainty calculated as follows:

$$U_f = \sqrt{(LOD/2)^2 + (\alpha C)^2}$$

where U_f is the maximum standard measurement uncertainty; LOD is the limit of detection of the method; C is the concentration of interest and α is a numeric factor depending on the value of C.

When the target measurement uncertainty is not defined, but guidance is given regarding other method performance characteristics, typically limit of detection, precision and recovery, a target measurement uncertainty can be derived from such information.

Alternatively, values for the target measurement uncertainty may be derived from the criteria applied by PT providers to evaluate the participants’ performance (the “standard deviation for proficiency assessment”) if they reflect acceptable performance for the intended use of the measurement requested to the laboratory. The results of collaborative studies or other interlaboratory comparisons can also be used in a similar way.

In some cases, the best option may be to adopt the same approach or values defined for other cases, which are technically similar, or related, to the one at hands.

The suggestions indicated here are far from being comprehensive. Depending on the problem at hand and the specific issues involved, including potential risks and economic costs associated with the decisions to be taken on the basis of the measurement results, specific studies may be

necessary, addressing the overall benefits of using measurement procedures of higher metrological order, i.e. lowering the target measurement uncertainty.

3.2. Assessing the plausibility of measurement uncertainty estimates

Participation in proficiency testing provides not only an assessment of the performance of the laboratory but also sufficient information for the laboratory to assess its own claim regarding measurement uncertainty. Provided that the assigned values are known with sufficient reliability, the laboratory may confirm the plausibility of its claim of measurement uncertainty, by visual inspection that the interval surrounding the measurement result overlaps with the assigned value. More specifically, the PT provider may issue zeta-scores to the participants, calculated according to:

$$\zeta_i = \frac{x_i - x_{pt}}{\sqrt{u^2(x_i) + u^2(x_{pt})}}$$

where x_i is the participant result; x_{pt} is the assigned value; $u(x_i)$ is the standard uncertainty reported by the participant and $u(x_{pt})$ to the standard uncertainty of the assigned value. The zeta-score is evaluated as satisfactory between 0 and 2; questionable between 2 and 3 and unsatisfactory beyond 3, which give a numerical grading to the appropriateness of their claim of uncertainty.

In addition the revised ISO 13528 [8] points out to providers of proficiency testing schemes the need to alert their participants regarding the reliability of their uncertainty estimates, in particular, for:

$$u(x_i) \ll u(x_{pt}) \quad \text{or} \quad u(x_i) >> 1.5 s^*$$

where s^* is the robust standard deviation of the participants' results.

The implementation of these recommendations should provide laboratories with the means to assess critically whether their claims of uncertainty is reliable and fit for the intended use of the measurement result.

3. CONCLUSIONS

More than twenty years after the first issuing of guidance on the evaluation and expression of uncertainty in measurement, the concept of uncertainty is being increasingly used as a key parameter for assessing measurement quality and defining the level of confidence that can be placed in the decisions taken on the basis of measurement results.

A number of regulatory and guidance documents have been published addressing several issues in connection with the estimate and the use of measurement uncertainty and they are increasingly adopted in everyday practice.

Further developments and improvements are therefore expected over the next years, leading to a harmonised application of the concept of measurement uncertainty, improved understanding of the factors having effects on measurement results and more effective implementation of the legislation, e.g. in food analysis, requiring compliance to specification to be demonstrated using measurement results.

REFERENCES

- [1] ISO/IEC Guide 98:1993. *Guide to the expression of uncertainty in measurement (GUM)*, ISO, Geneva, Switzerland, 1993. Corrected and reprinted in 1995. Superseded by [2] and [3].
- [2] Joint Committee for Guides in Metrology. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM)*, JCGM 100, BIPM, 2008 (www.bipm.org).
- [3] ISO/IEC Guide 98-3:2008. *Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*, ISO, Geneva, Switzerland, 2008.
- [4] S.L.R. Ellison and A. Williams (Eds). Eurachem/CITAC guide: *Quantifying Uncertainty in Analytical Measurement*, Third edition, Eurachem, 2012, www.eurachem.org
- [5] B. Magnusson, T. Näykki, H. Hovind and M. Krysell. *Handbook for calculation of measurement uncertainty in environmental laboratories*, Nordtest Report TR 537 (ed. 3.1) 2012, www.nordtest.info.
- [6] ISO 21748:2010. *Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation*. ISO, Geneva, Switzerland, 2010.
- [7] R. Bettencourt da Silva and A. Williams (Eds). Eurachem/CITAC Guide: *Setting and Using Target Uncertainty in Chemical Measurement*, First edition, Eurachem, 2015, www.eurachem.org.

- [8] ISO 13528:2015. *Statistical methods for use in proficiency testing by interlaboratory comparison*, ISO, Geneva, Switzerland, 2015.
- [9] V. J. Barwick, E. Prichard (eds.), Eurachem Guide: *Terminology in analytical measurement – Introduction to VIM 3*, Eurachem, 2011, ISBN 978-0-948926-29-7, www.eurachem.org.
- [10] B. Magnusson and U. Örnemark (Eds.) Eurachem Guide: *The Fitness for purpose of analytical methods – A Laboratory guide to method validation and related topics*, 2nd ed., Eurachem, 2014, ISBN 978-91-87461-59-0, www.eurachem.org
- [11] ISO 5725:1994. *Accuracy (trueness and precision) of measurement methods and results*, ISO, Geneva, Switzerland, 1994.
- [12] Joint Committee for Guides in Metrology, *International vocabulary of metrology – basic and general concepts and associated terms (VIM)*, Third edition, 2008, re-issued with minor corrections, JCGM 200, BIPM, 2012 (www.bipm.org).
- [13] Commission Regulation (EC) No 401/2006 of 23 February 2006 laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs
- [14] Commission Regulation (EC) No 333/2007 of 28 March 2007 laying down the methods of sampling and analysis for the official control of the levels of lead, cadmium, mercury, inorganic tin, 3-MCPD and polycyclic aromatic hydrocarbons in foodstuffs.