Issues and challenges in applicability of measurement uncertainty estimation in medical laboratories

Andrea Padoan[1,2], Giorgia Antonelli[1,2], Ada Aita[1,2], Laura Sciacovelli[2], Mario Plebani[1,2]

[1]Department of Medicine—DIMED, University of Padova, Padova, Italy; [2]Department of Laboratory Medicine, University-Hospital of Padova, Padova, Italy
Correspondence to: Andrea Padoan. Department of Medicine—DIMED, University of Padova, via Giustiniani 2, 35128, Padova, Italy. Email: andrea.padoan@unipd.it.

Abstract: The ISO 15189 has endorsed the use of measurement uncertainty (MU) to estimate the quantity of measurement error of a given method, although the total error (TE) approach, originally postulated by Westgard *et al.* in 1974, is still widely used and appreciated in laboratory medicine. Both MU and TE present some struggles, especially regarding the estimation of both bias and imprecision. In particular, bias may be seen as a complicating factor because it needs specification of a reference quantity, which may be obtained by different sources. Different formulations of MU have been proposed by different guidelines so far, and each approach presents advantages or limitations. These methods are not comparable and there are considerable differences not only in the theoretic approach but also in MU formulation. Moreover, only few practical results have been reported in the literature, thus limiting comprehension and applicability of these methods, ultimately supporting the hypothesis that implementation of MU in medical laboratories remains in embryo. Additional research and developments in this area are hence advisable for establishing consensus on practical methods for medical laboratories.

Keywords: Measurement uncertainty (MU); ISO 15189; total error (TE); measurement error; bias; imprecision

The measurement uncertainty (MU) introduction in the medical laboratories

The concepts of MU and its estimation represent two quite new notions for medical laboratories since their use has only increased after introduction of ISO 15189 in 2003. The Bureau international des poids et mesures (BIPM), an intergovernmental organization established for maintaining the international system unit since the late 1875, has endorsed the use of MU, firstly in the field of physics and, afterwards, in chemistry in the 1980s. Since the 1993, the BIPM has developed the document "Evaluation of measurement data—Guide to the expression of uncertainty in measurement" (GUM), aimed to provide a framework for evaluating MU (1). In particular, the basic concepts associated with MU and its definition are clearly reported in GUM (JCGM 100:2008), whilst the document also includes

a part identifying the possible source that should be considered for estimating the uncertainty of a measurement, along with practical consideration (2).

The MU is based on the observation that there cannot be measurement without uncertainty. Notably, a "measurement error" should not be considered a mistake but, more appropriately, a variability that is an inherent part of a measurement and of the measurement process itself. For a specific singleton measure, the measurement error can be considered the "measured quantity value, minus a reference quantity value". However, when considering a series of repeated measures, measurement error may be divided in two leading components, i.e., random and systematic errors. Random error represents the component of measurement error varying in an unpredictable manner, whilst systematic error does not vary (remains constant) or can change in a predictable manner (3).

A different concept is expressed by total error (TE), introduced by Westgard *et al.* in 1974. TE represents the expression of the total deviation of test result from its "true value" and, as originally formulated (TE = bias + Z × CV$_A$), TE includes both random and systematic error components. The inclusion of the statistical (unidirectional) Z value indicates that TE is calculated considering a confidence interval (typically as high as 95%) on the measurable error, despite a residual chance of 5% remains that a test result may exceed this limit. TE, thus, requires that the true value of an estimate is known, otherwise TE cannot be calculated (4). Conversely, MU assumes that the true value of a measurand cannot be exactly known.

According to these premises, two types of measurement error models can be derived: the TE model for "error methods", and the MU model for "uncertainty methods". Error methods require that a true value of a quantity is known, because error is an inherent property of measurements. Unlike errors methods, uncertainty methods do underlie the lack of a true value, but they claim the absence of an exact knowledge of the true value (4).

Issues on imprecision and bias estimation

Imprecision and bias are the two components to be considered in measurement error, as previously discussed. Imprecision represents the quantitative estimate of precision, a concept of quality defined by the international vocabulary of metrology (JCGM 200:2008, VIM) as the "*closeness of agreement between indications or measured quantity values, obtained by replicate measurements on the same of similar objects under specified conditions*". Imprecision, therefore, represents the numerical estimate of precision, and can be expressed as multiply of standard deviation, variance or coefficient of variation. Analogously, bias represents the quantitative numerical estimate of the degree of trueness, the latter defined by VIM as "*the closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value*". Imprecision and bias are thus two numeric expressions for random and systematic error, respectively, and both quantities are mutually exclusive (5).

Different definitions of imprecision currently exist, depending on the measurement condition. For estimation of repeatability, the imprecision is estimated by a set of conditions including the same laboratory, the same equipment, the same material and the same staff, within a short time frame. Intermediate imprecision include, again,

that measurements are performed in the same lab, but with changes to calibrators and operators; reproducibility should be estimated by including long-term variability and changes of measurement methods and laboratories (5).

Regardless of the type of measurement error methods (error or uncertainty methods), estimation of bias, differently from imprecision, represents a more challenging concept. Even if bias definition is clear and understandable, its estimation includes the use of a reference quantity value. Reference quantities may be obtained either by certified reference materials (CRM) or by reference measurement procedures (RMP), both guarantying metrological traceability of data (6). However, CRM and RMP are only available for a limited number of measurands, thus limiting the applicability of the calculation of bias through reference quantities. Alternatively, assigned values derived by a consensus agreement, either by external quality assessment schemes (EQAs) or proficiency testing (PT) may be used (3). In this case, the bias estimation presents limitations, not only because the procedure used by EQAs or PT organizers for establishing the assigned value may change depending on the statistical approach, but also because different assigned values may be derived for all laboratories in the survey or for specified subgroups (e.g., based on measuring systems, analytical methods, etc.) (7).

The MU estimation approaches need harmonization

Several approaches for estimating MU for medical laboratories have been suggested so far. Nevertheless, the basic definition of MU, "*a non-negative parameter characterizing the dispersion of the quantity value being attributed to a measurand, based on the information used*" is well consolidated and shared among the developed MU guidelines (5).

The GUM guideline proposed two methods for evaluating the MU, i.e., the Type A and Type B evaluations. A coverage factor (k) is also used as multiplier of combined standard uncertainty for obtaining expanded uncertainty, the quantity usually referred with MU (2). The Eurachem/CITAC Guide CG 4 2012 guideline used a similar approach, specifying that all sources of uncertainty should be measured and included in MU estimation (8). The Clinical and Laboratory Standard Institute (CLSI) EP29-A guidelines has proposed two approached, namely the

"top-down" and the "bottom-up". The former approach can be used to estimate the MU from the long-term QC data routinely collected in most laboratories, whilst the latter approach can be used when measurement procedure performance data (collected in verification experiments or from information provided by manufacturers) can be available. Although all the three guidelines consider imprecision as the main component of MU, bias, if not negligible, should be removed by either applying a correction to measurements or through instrumental recalibration (9). Interestingly, the term "negligible bias" is not clearly defined, though it should be either modest or clinical insignificant. When the bias cannot be eliminated, it can be handled as any other source of uncertainty and thus included in calculation of MU, as suggested by Theodorsson (4).

The Nordtest guideline presents a different approach for MU estimation. The bias, expressed as root mean square, is expected to be included in the calculation, along with within-laboratory reproducibility (intermediate imprecision). The guideline also showed different methods for estimating bias that could be calculated by CRM, by interlaboratory comparisons (e.g., EQAs or PT), or by recovery experiments (10). Therefore, the Nordtest approach appears flexible and suitable to several assays. We have evaluated the practical applicability of MU estimation for a total of 263 tests, by adapting the Nordtest-based method, observing that the derived approach could be effectively applied to a large series of measurands, using data already available in the laboratory, such as long-term quality control material data and EQAs results (11).

Recently, it has been proposed that MU could include other sources of uncertainty from the total testing process (TTP), especially the biological pre-analytical and post-analytical variability, when these can be accurately estimated (4,12,13). However, the ISO 15189:2012 standard, in chapter 5.5.1.4, mentions that *"The relevant uncertainty components are those associated with the actual measurement process, commencing with the presentation of the sample to the measurement procedure and ending with the output of the measured value"*. Therefore, the inclusion of pre-analytical and biological variability is not requested in MU, at least for purposes of ISO 15189:2012 accreditation (14).

Challenges to the practical application of the MU estimation

Regardless of the approach chosen for MU estimation, some

issues remained unresolved and still represent challenges for pragmatic applicability of the proposed theoretical principles.

The first issue to be addressed concerns the number of levels to be considered for MU estimation, for each measurement procedure. The degree of association between the measurand concentration levels, along with imprecision and bias, should be carefully inspected and two or more MU values should be estimated for measurement procedures where bias, imprecision (or both) are expectedly dependent on concentration levels.

A second important aspect concerns the fit-for-purpose of test results. When a test results is used for patient monitoring, the imprecision component should only be included in MU estimation. As a reliable example, two consecutive results with non-negligible bias characterized by similar biases in both measurements, they will finally elide each other. At variance, bias inclusion is appropriate when test results are related to a clinical decision point (15).

Conclusions

The mission of medical laboratories is to guarantee the quality of patients' results, in order for improving clinicians' understanding of diseases conditions or monitoring treatments. To fulfil this aim, laboratory should estimate and monitor the measurement error of their analytical procedures. Currently, error model based on TE are widely used, since they include imprecision and bias components effecting results of analytical measurement. Nevertheless, the ISO 15189 endorsed the usage of uncertainty error model, based on estimation of MU. Therefore, medical laboratories are supposed to estimate MU for all measurement procedures undergoing or subjected to ISO 15189 accreditation.

Currently, the MU methods in laboratory medicine faces several challenges, such as (I) need of harmonization of proposed approaches; (II) limitations of bias estimation and (III) unfamiliarity with statistical calculation needed for MU, combined with scarce availability of practical examples in literature. Despite these important caveats, which are probably related to the embryonic phase of implementation of MU in medical laboratory, the estimation of MU for ISO 15189 accreditation may be seen as an opportunity rather than as a change (with respect to TE) for clinical laboratories and for overall quality assurance of test results.

Acknowledgments

Footnote

Provenance and Peer Review: This article was commissioned by the editorial office, *Journal of Laboratory and Precision Medicine* for the series "International Conference on Laboratory Medicine". The article has undergone external peer review.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jlpm.2017.08.10). The series "International Conference on Laboratory Medicine" was commissioned by the editorial office without any funding or sponsorship. Mario Plebani served as an unpaid Guest Editor of the series. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Farrance I, Badrick T, Sikaris KA. Uncertainty in measurement and total error - are they so incompatible? Clin Chem Lab Med 2016;54:1309-11.
2. JCGM 100:2012. Evaluation of measurement data - guide to the expression of uncertainty in measurement (GUM). Available online: www.bipm.org
3. Oosterhuis WP, Bayat H, Armbruster D, et al. The use of error and uncertainty methods in the medical laboratory. Clin Chem Lab Med 2017. [Epub ahead of print].
4. Theodorsson E. Uncertainty in Measurement and Total Error: Tools for Coping with Diagnostic Uncertainty. Clin Lab Med 2017;37:15-34.
5. JCGM 200:2012. International Vocabulary of Metrology – Basic and General Concepts and Associated Terms. (VIM 3rd edition). Available online: http://www.bipm.org/en/publications/guides/vim.html
6. Jones GR, Jackson C. The Joint Committee for Traceability in Laboratory Medicine (JCTLM) - its history and operation. Clin Chim Acta 2016;453:86-94.
7. Jones GRD, Albarede S, Kesseler D, et al. Analytical performance specifications for external quality assessment - definitions and descriptions. Clin Chem Lab Med 2017;55:949-55.
8. Quantifying uncertainty in analytical measurement. Available online: https://www.eurachem.org/index.php/publications/guides/quam
9. Clinical and Laboratory Standards Institute (CLSI). Expression of measurement uncertainty in Laboratory Medicine. Approved Guideline. 2012. Available online: https://clsi.org/media/1420/ep29a_sample.pdf
10. Magnusson B, Näykk T, Hovind H, et al. Handbook for calculation of Measurement Uncertainty in Environmental Laboratories. NT TR 537 - Edition 3.1, 2012. Available online: http://www.nordtest.info/index.php/technical-reports/item/handbook-for-calculation-of-measurement-uncertainty-in-environmental-laboratories-nt-tr-537-edition-3.html
11. Padoan A, Antonelli G, Aita A, et al. An approach for estimating measurement uncertainty in medical laboratories using data from long-term quality control and external quality assessment schemes. Clin Chem Lab Med 2017. [Epub ahead of print].
12. Carobene A, Røraas T, Sølvik UØ, et al. Biological Variation Estimates Obtained from 91 Healthy Study Participants for 9 Enzymes in Serum. Clin Chem 2017;63:1141-50.
13. Carobene A, Strollo M, Jonker N, et al. Sample collections from healthy volunteers for biological variation estimates' update: a new project undertaken by the Working Group on Biological Variation established by the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2016;54:1599-608.
14. ISO 15189. Medical laboratories—requirements for quality and competence. Geneva: International Organization for

Standardization (ISO), 2012. Available online: https://www.iso.org/obp/ui/#iso:std:56115:en

15. Tate JR, Plebani M. Measurement uncertainty - a revised understanding of its calculation and use. Clin Chem Lab Med 2016;54:1277-9.