

## Mini Review

Federica Braga\* and Mauro Panteghini

# The utility of measurement uncertainty in medical laboratories

<https://doi.org/10.1515/cclm-2019-1336>

Received December 30, 2019; accepted January 31, 2020

**Abstract:** The definition and enforcement of reference measurement systems, based on the implementation of metrological traceability of patient results to higher-order (reference) methods and/or materials, together with a clinically acceptable level of measurement uncertainty (MU), are fundamental requirements to produce accurate and equivalent laboratory results. The MU associated with each step of the traceability chain should be governed to obtain a final combined MU on clinical samples fulfilling the requested performance specifications. MU is useful for a number of reasons: (a) for giving objective information about the quality of individual laboratory performance; (b) for serving as a management tool for the medical laboratory and *in vitro* diagnostics (IVD) manufacturers, forcing them to investigate and eventually fix the identified problems; (c) for helping those manufacturers that produce superior products and measuring systems to demonstrate the superiority of those products; (d) for identifying analytes that need analytical improvement for their clinical use and ask IVD manufacturers to work for improving the quality of assay performance and (e) for abandoning assays with demonstrated insufficient quality. Accordingly, the MU should not be considered a parameter to be calculated by medical laboratories just to fulfill accreditation standards, but it must become a key quality indicator to describe both the performance of an IVD measuring system and the laboratory itself.

**Keywords:** measurement uncertainty; metrological traceability; performance specifications; standardization.

---

\*Corresponding author: Federica Braga, Research Centre for Metrological Traceability in Laboratory Medicine (CIRME), Università di Milano, Milan, Italy; and UOC Patologia Clinica, ASST Fatebenefratelli-Sacco, Via GB Grassi 74, 20157 Milan, Italy, Phone: +390239042743, Fax: +390250319835, E-mail: federica.braga@unimi.it

Mauro Panteghini: Research Centre for Metrological Traceability in Laboratory Medicine (CIRME), Università di Milano, Milan, Italy

## Introduction

In the International Vocabulary of Metrology, the measurement uncertainty (MU) is defined as a “parameter characterizing the dispersion of the quantity values being attributed to a measurand” [1]. It describes the interval within which the value of the measurand is assumed to lie with a stated level of confidence. If in the general use, the term “uncertainty” relates to the concept of doubt, in medical laboratory the knowledge of MU implies instead increased confidence in the validity of a measurement result [2, 3].

The knowledge of MU and the definition of its allowable limits for the clinical application of measurements represents one of the mainstays, together with the definition of reference measurement systems and the establishment of a proper post-market surveillance of *in vitro* diagnostics (IVD) quality, needed to produce standardized laboratory results suitable for clinical use [3]. Estimating and checking MU in medical laboratories is essential to understand its influence on measurement results and, ultimately, their clinical suitability. However, although all medical laboratories seeking ISO 15189 accreditation know that MU estimate is a specific requirement (clause 5.5.1.4), few know what to do with the calculated MU [4].

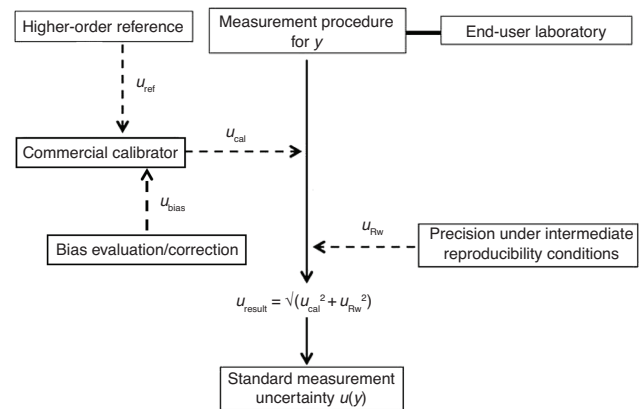
## How to calculate MU in medical laboratories

Historically, two approaches have been proposed for estimating MU: the so-called “bottom-up” and “top-down” approaches [5]. The “bottom-up” approach is the model originally proposed by the “Guide to the Expression of Uncertainty of Measurement” (GUM) [6]. This model, usually employed by reference laboratories to obtain accreditation according to ISO 17025 and 15195 standards, is based on a comprehensive dissection of the measurement, in which each potential source of uncertainty is identified, quantified and then combined to generate the MU of the result using statistical propagation rules [2]. We previously described the application of

this approach to the enzyme measurements using IFCC reference measurement procedures [7]. The application of this approach in medical laboratories is however too complicated and has encountered many practical problems and objections [8].

The “top-down” approach is simpler and represents a good alternative to the previous approach. It estimates MU of laboratory results by using internal quality control (IQC) data to derive the random components of uncertainty and commercial calibrator information. It is now officially endorsed by the ISO Technical Specification 20914 that provides a practical guidance to be applied in medical laboratory settings for the purpose of estimating MU of values produced by measurement procedures intended to measure biological measurands [9]. The inspiring concept behind this approach, described in Figure 1, relies on the definition of MU across the entire traceability chain, starting with the uncertainty of reference materials ( $u_{\text{ref}}$ ), extending through the IVD manufacturers and their processes for assignment of calibrator values and uncertainty ( $u_{\text{cal}}$ ) and ending with the random variability of measuring systems ( $u_{\text{Rw}}$ )<sup>1</sup> [10]. In particular,  $u_{\text{Rw}}$  can be derived from IQC, while  $u_{\text{cal}}$  must include all uncertainties introduced by the selected calibration hierarchy for the measurand beginning with the highest available reference down to the assigned value of the calibrator for the commercial IVD medical device, including the uncertainty of bias correction ( $u_{\text{bias}}$ ), if a not negligible bias has been detected and corrected by the manufacturer when implementing traceability.

Although reference material providers, IVD manufacturers and medical laboratories have different roles and independent tasks across the metrological traceability chain, their performances contribute together to the MU of patient results [11]. The crucial point is that the estimated MU must be always combined at each level of the employed traceability chain. Particularly, the MU at the level of clinical samples ( $u_{\text{result}}$ ) must be the combination of all uncertainty contributions accumulated across the entire traceability chain ( $u_{\text{result}} = \sqrt{u_{\text{cal}}^2 + u_{\text{Rw}}^2}$ ), where  $u_{\text{cal}} = \sqrt{u_{\text{ref}}^2 + u_{\text{value assignment}}^2 + u_{\text{bias}}^2}$ , if any). This refutes the common misconception that the simple reproducibility of a measurement result equals its overall MU. A correct estimate of MU of laboratory results is indeed not possible without  $u_{\text{cal}}$ . In the European market, the information



**Figure 1:** Sources of measurement uncertainty across the entire metrological traceability chain.

about  $u_{\text{cal}}$  shall be provided on request to the professional end-users. Sometimes, calibrators are offered without uncertainty, but it is up to the laboratory professionals to pretend this information and, in case of unavailability, the corresponding material should be disregarded and replaced with some alternatives offering this information needed for the correct estimate of  $u_{\text{result}}$ . On the other hand, it is very important to define conditions for deriving  $u_{\text{Rw}}$  that should correspond to a within-laboratory reproducibility for a period (e.g. 6 consecutive months) sufficient to capture most changes to measuring conditions and systematic sources of uncertainty, such as those caused by different lots of reagents, different calibrations or different environmental conditions [9]. Characteristics of control material for estimating  $u_{\text{Rw}}$  have been defined and should be carefully considered, i.e. the material should be different from that used to check the correct alignment of the measuring system, be commutable and with concentration(s) corresponding to the decision cut-point(s) employed in the medical application of the test [12].

We previously reported some practical examples on how medical laboratories can correctly calculate  $u_{\text{result}}$ . As an example, figure 1 of ref. [13] described the metrological traceability chain and combined standard MU of Abbott Architect creatinine enzymatic assay. In particular,  $u_{\text{Rw}}$  was estimated as CV from 6-month consecutive measurement data of a serum-based fresh-frozen control material, randomly analyzed daily during our ordinary laboratory activity.

## How to define maximum allowable MU

The ISO Technical Specification 20914 is also clear in pointing out that the magnitude of MU should be suitable

<sup>1</sup> ISO/TS 20914:2019 defines  $u_{\text{Rw}}$  as “uncertainty component under conditions of within-laboratory precision” (i.e. the uncertainty for a given measuring system in the same laboratory over an extended time period that includes routine changes to measuring conditions, for example, lot changes of reagents, calibrators, instrument maintenance, etc.).

for a result to be used in a medical decision: “for a given measuring system, estimating the uncertainty of the results produced is of very limited value unless it can be compared with the allowable MU based on the quality of results required for medical use” [9]. After the 1st European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Strategic Conference, held in 2014, objective criteria for defining analytical performance specifications (APS) became available. These criteria are based on three models: model 1, based on the effect of analytical performance on clinical outcomes; model 2, based on components of biological variation of the measurand; and model 3, based on state of the art of the measurement (defined as the highest level of analytical performance technically achievable) [14]. One revolutionary aspect of this approach was to emphasize that certain models are better suited for some measurands than for others, and the attention should therefore primarily direct toward the measurand and its biological and clinical characteristics [15]. Grading different levels of quality (i.e. minimum, desirable and optimum) for APS is also very important because it stimulates IVD manufacturers to work for improving the quality of assays to move, in case, from unacceptable or minimum to desirable performance [16].

For MU, the relevant goal that should be fulfilled is that related to the allowable random variability of patient results, as the correct trueness transfer along the metrological traceability chain should allow the achievement of unbiased (or negligibly biased) results [17]. In a recent paper, we used serum creatinine as an example for the definition of maximum allowable MU [18]. This measurand has a strict metabolic control so that the most appropriate model for deriving APS is that based on its biological variation. By using published data about the average intra-individual biological variation ( $CV_i$ ) of serum creatinine (4.4%) [19] and the classical Fraser’s paradigm for deriving APS for random variability [20], APS for standard MU of serum creatinine measurement on clinical samples are 3.3% ( $\leq 0.75 CV_i$ , minimum quality), 2.2% ( $\leq 0.50 CV_i$ , desirable quality) and 1.1% ( $\leq 0.25 CV_i$ , optimum quality).

Once defined, APS cover the total MU budget ( $TB_u$ ) that should be fulfilled at the level of patient results. The achievement of  $TB_u$  depends on the MU contributions of each step of the metrological traceability chain [11, 12]. Therefore, it is essential to accurately define the entity of all those contributions and how much of the  $TB_u$  is used across the different steps of traceability chain. We previously recommended that specific MU limits at different levels of the traceability chain should be defined as fractions of allowed  $TB_u$ ; in particular, we conventionally recommended that no more than one third of  $TB_u$  should

be consumed by  $u_{ref}$  and  $\leq 50\%$  of  $TB_u$  used by  $u_{cal}$ . The remaining MU should be available for  $u_{Rw}$  as a margin to fulfill  $TB_u$  [11, 12].

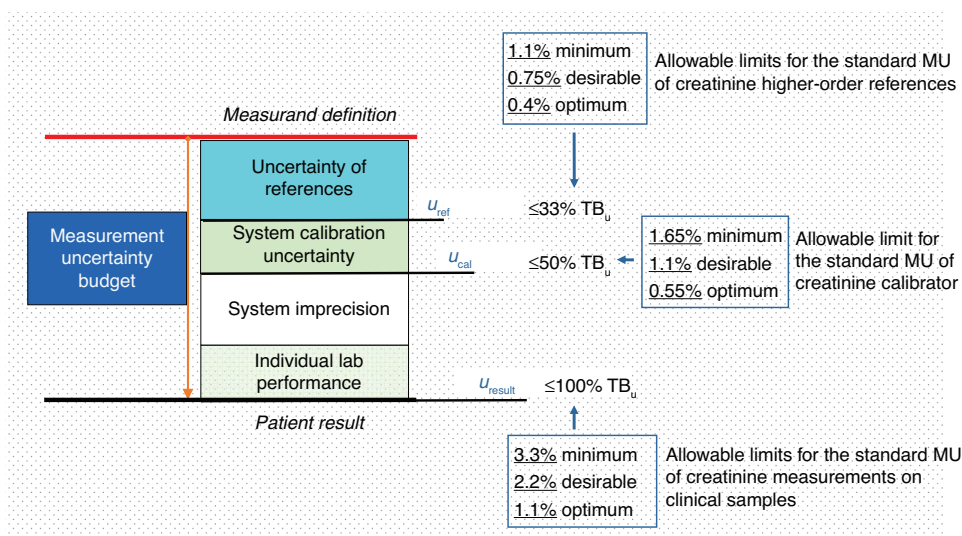
The  $u_{ref}$  represents the first contribution to the  $TB_u$ . Due to uncertainty propagation in the calibration hierarchy,  $u_{ref}$  may significantly affect the MU of patient results. It is therefore intuitive that it should be markedly lower than APS for  $TB_u$ . Continuing to use serum creatinine as an example, Figure 2 reports how to derive the allowable limits for the standard MU of higher-order references in order to not exceed with a high probability  $TB_u$  at the level of clinical samples [18].

The role of IVD manufacturers is to identify higher-order metrological references and, based on them, to define a calibration hierarchy to assign traceable values to system calibrators and estimate their MU [21]. The basic paradigm here is that, if present in a not negligible amount, a systematic error (bias) should be appropriately eliminated by adjusting the value assigned to the calibrator, while the overall MU increases because of the  $u_{bias}$  contribution [22]. In addition to  $u_{bias}$ , the manufacturer must also combine  $u_{ref}$ . Once estimated,  $u_{cal}$  should be compared with MU limits, which represent a proportion, e.g. 50%, of the  $TB_u$  allowed for clinical laboratory results (Figure 2). Keeping  $u_{cal}$  to a level fulfilling clinical needs is however a highly disregarded issue [13, 23–25]. When higher-order references do not exist, commercial calibrators are usually value-assigned by manufacturers using in-house procedures. However, even in this case, end-user calibrator assigned values will have an MU that contributes to the  $u_{result}$ . In these circumstances,  $u_{cal}$  will simply correspond to  $u_{value\ assignment}$  (see description in the section “How to calculate MU in medical laboratories”), with no contribution from  $u_{ref}$  and  $u_{bias}$ .

In a previous paper, we provided simulations about the status of the uncertainty budget for some measurands [11]. The indicated approach should now be applied to each analyte measured in the medical laboratory to verify if the status of the uncertainty budget of its measurement associated with the selected metrological traceability chain is suitable for clinical application of the test.

## How to deal with bias on clinical measurements

In the traceability framework, medical laboratories should rely on the IVD manufacturers who must ensure traceability of their measuring systems to the highest available references. Therefore, regular estimation of bias by the



**Figure 2:** Allowable limits for the standard measurement uncertainty (MU) of serum creatinine results on clinical samples (derived from the biological variation model) ( $u_{result}$ ) and corresponding limits for creatinine higher-order references (materials or procedures) ( $u_{ref}$ ) and commercial calibrator ( $u_{cal}$ ), expressed as a fraction of the total uncertainty budget ( $TB_u$ ).

end-user laboratory is not required. As the IVD measuring system is CE (“Communautés Européennes”)-marked and correct alignment to higher-order references is expected, laboratories should just consider the MU of the value assigned to the calibrator (that should include the  $u_{bias}$ , if any) and combine it with  $u_{RW}$  to obtain  $u_{result}$ . Appearance of a medically unacceptable measurement bias could be however shown by external quality assessment (EQA) surveillance, but caution needs to be exercised as only schemes fulfilling category I/IIA criteria are usable to this scope [13, 26]. If a medically significant bias is suspected during ongoing EQA surveillance, the bias against a reference (material or procedure) for that measurand should be estimated and the presence of a significant systematic error confirmed. Then, the bias value should be included in the estimate of MU of clinical samples [23, 27]. If the recalculated MU is not fulfilling the predefined APS, it is the responsibility of the manufacturer to take an immediate investigation and eventually fix the problem with a corrective action (e.g. by improving the calibrator value-assignment protocol). If unsolved by the manufacturer, the laboratory could introduce a correction factor for the detected bias. If so, the uncertainty of the correction factor needs to be estimated and included in the calculation of  $u_{result}$ . The use of bias correction factors by individual laboratories is however not permitted by some national regulations as this may alter the status of the measuring system, removing any responsibility from the manufacturer and depriving the system (and, consequently, the produced results) of the certification originally provided through

CE marking. The introduction of correction factors by individual laboratories is also quite risky, as they are usually unaware of possible subsequent changes made by the manufacturer and may continue to use the correction factor even when the bias has been corrected in the reagent production stage. Therefore, we are not supporting the individual use of bias correction factors in daily practice, but we strongly believe that involved laboratories should insist in order that the providing manufacturer quickly solves the issue. Table 1 summarizes the suggested sequential approach in case a clinically significant bias on patient results is suspected.

## Why MU matters in medical laboratories

### Suitability and selection of higher-order references

As discussed earlier, MU of higher-order references may significantly influence the fulfillment of APS for  $u_{result}$ . Using plasma glucose as an example, we demonstrated that at least four different metrological traceability chains can be used to transfer trueness from the measurand definition (according to the International System of Units [SI]) to commercial calibrators. By selecting one of these chains, IVD manufacturers may spend however very different amounts of the  $TB_u$  in implementing traceability of

**Table 1:** Steps related to how to deal with bias on clinical measurements.

- 
1. Discover a medically unacceptable measurement bias during the external quality assessment (EQA) program (only schemes fulfilling category I/IIA criteria are however usable to this scope)
  2. If a medically significant bias (meaning a bias that does not fulfill the corresponding performance specifications) is suspected during the ongoing EQA surveillance, the bias against a reference (material or procedure) for that measurand should be estimated and the presence of a significant systematic error confirmed. Note that as reference may act any material or procedure positioned at the top of the corresponding traceability chain, even in the absence of high-order options
  3. The obtained bias value should be included in the estimate of measurement uncertainty (MU) of clinical samples
  4. If the recalculated MU is not fulfilling the predefined performance specifications, it is the responsibility of the manufacturer to take an immediate investigation and eventually fix the problem with a corrective action
  5. If unsolved by the manufacturer, the individual laboratory could introduce a correction factor for the detected bias. If so, the uncertainty of the correction factor needs to be estimated and included in the calculation of  $u_{\text{result}}$ . Note that the use of bias correction factors by individual laboratories may significantly alter the status of the commercial measuring system, removing the manufacturer's responsibility and depriving the system of the certification originally provided through CE marking
- 

their measuring systems [21, 28]. Therefore, the quality of glucose measurement may be dependent on the type of traceability chain selected by manufacturers for trueness transferring, sometimes making it difficult to achieve APS for MU at the level of clinical samples. This is also true for other measurands, like serum creatinine [12, 18]. Accordingly, when different options are available, in making choice IVD manufacturers should start to consider the suitability of higher-order references in terms of  $u_{\text{ref}}$  by selecting that with less impact on  $TB_u$  [18].

## Verification of quality of IVD medical devices

$u_{\text{cal}}$  may significantly impact the quality of IVD medical devices. However, in a number of studies we have shown that the manufacturer's internal quality specifications to validate the calibrator traceability to higher-order references and then derive  $u_{\text{cal}}$  are more often not established on the basis of suitable APS [23–25]. For example, Abbott Diagnostics in a document released in August 2014 informed customers that the internal release specification for serum creatinine calibrators was  $\pm 5\%$  from the target value of National Institute of Standards and Technology (NIST) SRM 967a Level 1 [23]. However, this clearly does not

fit with the aforementioned desirable APS for standard MU of creatinine measurements on clinical samples ( $\pm 2.2\%$ ). Similarly, for serum total folate, Beckman Coulter in a technical bulletin released in 2011 informed customers that the internal release specification for their calibrators was  $\pm 10\%$  from the target value of WHO International Standard 03/178 [25]. Once again, this does not fit with APS for standard MU of folate on clinical samples, which, for assuring a clinically acceptable misclassification rate of individuals with suspected vitamin deficiency, should remain within  $\pm 2.5\%$  [29]. Manufacturers should therefore start to conform their internal protocols of trueness transfer from certified reference materials to commercial calibrators to the clinical value of tests and their APS defined according to the aforementioned models.

## Providing evidence of clinically unsuitable results and stimulate work for improving the quality of assay performance

We previously reported examples in which deriving MU of clinical results provided evidence of clinically unsuitable results and stimulate manufacturers to work for improving the quality of assay performance [3]. In a first case, the performance of the immunoturbidimetric assay by Roche Diagnostics for measuring serum albumin, originally showing a too large bias and consequently an MU unable to fulfill APS for the clinical use of the test [27], was significantly improved in its metrological alignment resulting in the best recovery of the ERM-DA470k/IFCC reference material among 24 commercial assays for serum albumin [30]. Similarly, the status of glycated hemoglobin ( $HbA_{1c}$ ) traceability and associated MU, judged not enough in a detailed analysis [31], was later significantly improved permitting to fulfill APS for MU of  $HbA_{1c}$  results [32].

## Conclusions

In this paper, we provided an overview on how MU should be correctly estimated by medical laboratories and about the importance of this activity to guarantee reliability and quality of provided results. These concepts should apply to both marketed measuring systems and in-house procedures, when CE-marked commercial alternatives are not available. Medical laboratories should estimate (by using the “top-down” approach) and validate (by suitable APS) the MU of each test at the level of patient results. This activity is useful for several reasons, summarized in Table 2, all aimed at improving the quality of patient results provided

**Table 2:** Reasons supporting the measurement uncertainty estimate in medical laboratories.

- 
- Giving objective information about the quality of individual laboratory performance, supporting in turn a proper clinical decision-making process
  - Serving as a management tool for the clinical laboratory and IVD manufacturers, forcing them to investigate and eventually fix the identified problem
  - Helping those manufacturers that produce superior products and measuring systems to demonstrate the superiority of those products
  - Identifying measurands that need analytical improvement for their clinical use and ask IVD manufacturers to work for improving the quality of assay performance
  - Abandonment of assays with demonstrated insufficient quality
- 

by medical laboratories and, ultimately, the patient safety [33]. Considering that MU should be used to evaluate both the performance of an IVD measuring system and the laboratory itself, it should be added to the list of key quality indicators in all laboratories [34]. If needed, all attempts must be made to improve critical situations and, if the MU cannot be sufficiently reduced in order to fulfill APS, a decision can be made as to whether the measurement procedure is to be replaced with another performing better.

**Author contributions:** All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

**Research funding:** None declared.

**Employment or leadership:** None declared.

**Honorarium:** None declared.

**Competing interests:** The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

## References

1. JCGM 200:2012. International vocabulary of metrology – basic and general concepts and associated terms (VIM), 3rd ed. [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_200\\_2012.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2012.pdf). Accessed: Feb 2020.
2. Ellison SL, Williams A. Eurachem guide: quantifying uncertainty in analytical measurement. Eurachem, 3rd ed. 2012. [https://www.eurachem.org/images/stories/Guides/pdf/QUAM2012\\_P1.pdf](https://www.eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf). Accessed: Feb 2020.
3. Infusino I, Panteghini M. Measurement uncertainty: friend or foe? *Clin Biochem* 2018;57:3–6.
4. Westgard SA. Rhetoric versus reality? Laboratory surveys show actual practice differs considerably from proposed models and mandated calculations. *Clin Lab Med* 2017;37:35–45.
5. Topic E, Nikolac N, Panteghini M, Theodorsson E, Salvagno GL, Miler M, et al. How to assess the quality of your analytical method? *Clin Chem Lab Med* 2015;53:1707–18.
6. JCGM 100:2008. Evaluation of measurement data – guide to the expression of uncertainty in measurement (GUM). [https://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](https://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf). Accessed: Feb 2020.
7. Infusino I, Schumann G, Ceriotti F, Panteghini M. Standardization in clinical enzymology: a challenge for the theory of metrological traceability. *Clin Chem Lab Med* 2010;48:301–7.
8. Panteghini M. Application of traceability concepts to analytical quality control may reconcile total error with uncertainty of measurement. *Clin Chem Lab Med* 2010;48:7–10.
9. ISO/TS 20914:2019. Medical laboratories – practical guidance for the estimation of measurement uncertainty. 1st ed. Geneva, Switzerland: ISO, 2019.
10. Panteghini M. Implementation of standardization in clinical practice: not always an easy task. *Clin Chem Lab Med* 2012;50:1237–41.
11. Braga F, Panteghini M. Defining permissible limits for the combined uncertainty budget in the implementation of metrological traceability. *Clin Biochem* 2018;57:7–11.
12. Braga F, Infusino I, Panteghini M. Performance criteria for combined uncertainty budget in the implementation of metrological traceability. *Clin Chem Lab Med* 2015;53:905–12.
13. Braga F, Pasqualetti S, Panteghini M. The role of external quality assessment in the verification of in vitro medical diagnostics in the traceability era. *Clin Biochem* 2018;57:23–8.
14. Sandberg S, Fraser CG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: consensus statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med* 2015;53:833–5.
15. Ceriotti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, et al. Criteria for assigning laboratory measurands to models for analytical performance specifications defined in the 1st EFLM Strategic Conference. *Clin Chem Lab Med* 2017;55:189–94.
16. Panteghini M, Ceriotti F, Jones G, Oosterhuis W, Plebani M, Sandberg S, et al. Strategies to define performance specifications in laboratory medicine: 3 years on from the Milan Strategic Conference. *Clin Chem Lab Med* 2017;55:1849–56.
17. Bais R, Armbruster D, Jansen RT, Klee G, Panteghini M, Passarelli J, et al. Defining acceptable limits for the metrological traceability of specific measurands. *Clin Chem Lab Med* 2013;51:973–9.
18. Panteghini M, Braga F. Implementation of metrological traceability in laboratory medicine: where we are and what is missing. *Clin Chem Lab Med* 2020 Feb 18; doi: 10.1515/cclm-2019-1128. [Epub ahead of print]
19. Carobene A, Marino I, Coşkun A, Serteser M, Unsal I, Guerra E, et al. The EuBIVAS project: within- and between-subject biological variation data for serum creatinine using enzymatic and alkaline picrate methods and implications for monitoring. *Clin Chem* 2017;63:1527–36.
20. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem* 1997;34:8–12.
21. Braga F, Panteghini M. Verification of in vitro medical diagnostics (IVD) metrological traceability: responsibilities and strategies. *Clin Chim Acta* 2014;432:55–61.

22. Kallner A. Measurement performance goals: how they can be estimated and a view to managing them. *Scand J Clin Lab Investig* 2010;70(Suppl 242):34–9.
23. Pasqualetti S, Infusino I, Carnevale A, Szóke D, Panteghini M. The calibrator value assignment protocol of the Abbott enzymatic creatinine assay is inadequate for ensuring suitable quality of serum measurements. *Clin Chim Acta* 2015;450:125–6.
24. Braga F, Infusino I, Frusciante E, Ceriotti F, Panteghini M. Trueness evaluation and verification of interassay agreement of 11 serum IgA measuring systems: implications for medical decisions. *Clin Chem* 2019;65:473–83.
25. Braga F, Frusciante E, Ferraro S, Panteghini M. Trueness evaluation and verification of inter-assay agreement of serum folate measuring systems [published online ahead of print, 2020 Jan 11]. *Clin Chem Lab Med* 2020 Jan 11; doi: 10.1515/cclm-2019-0928. [Epub ahead of print].
26. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/external quality assessment: current challenges and future directions. *Clin Chem* 2011;57:1670–80.
27. Infusino I, Braga F, Mozzi F, Valente C, Panteghini M. Is the accuracy of serum albumin measurements suitable for clinical application of the test? *Clin Chim Acta* 2011;412:791–2.
28. Pasqualetti S, Braga F, Panteghini M. Pre-analytical and analytical aspects affecting clinical reliability of plasma glucose results. *Clin Biochem* 2017;50:587–94.
29. Ferraro S, Lyon AW, Braga F, Panteghini M. Definition of analytical quality specifications for serum total folate measurements using a simulation outcome-based model. *Clin Chem Lab Med* 2020;58:e66–8.
30. Bachmann LM, Yu M, Boyd JC, Bruns DE, Miller WG. State of harmonization of 24 serum albumin measurement procedures and implications for medical decisions. *Clin Chem* 2017;63:770–9.
31. Braga F, Panteghini M. Standardization and analytical goals for glycated hemoglobin measurement. *Clin Chem Lab Med* 2013;51:1719–26.
32. Szóke D, Carnevale A, Pasqualetti S, Braga F, Paleari R, Panteghini M. More on the accuracy of the Architect enzymatic assay for hemoglobin A<sub>1c</sub> and its traceability to the IFCC reference system. *Clin Chem Lab Med* 2016;54:e71–3.
33. Plebani M, Sciacovelli L, Bernardi D, Aita A, Antonelli G, Padoan A. What information on measurement uncertainty should be communicated to clinicians, and how? *Clin Biochem* 2018;57:18–22.
34. Sciacovelli L, Panteghini M, Lippi G, Sumarac Z, Cadamuro J, De Olivera Galoro CA, et al. Defining a roadmap for harmonizing quality indicators in Laboratory Medicine: a consensus statement on behalf of the IFCC Working Group “Laboratory Error and Patient Safety” and EFLM Task and Finish Group “Performance specifications for the extra-analytical phases”. *Clin Chem Lab Med* 2017;55:1478–88.