**Editorial**

Mauro Panteghini and Sverre Sandberg

# Total error vs. measurement uncertainty: the match continues

DOI 10.1515/cclm-2015-1036

Although 5 years have passed since the publication of a previous editorial in this journal that tried to reconcile the concepts [1], laboratory professionals are still playing the total error (TE) vs. uncertainty game. Two papers published in this issue of *Clinical Chemistry and Laboratory Medicine* further discuss the subject [2, 3], which was revitalized by the organization of the 1st EFLM Strategic Conference in Milan on the analytical performance specifications at the end of 2014 [4]. The organizers considered it timely to address the topic of performance specifications both because it was 15 years since it was previously addressed and because performance specifications are more and more central for the clinical application of test measurements [5, 6].

Being members of the scientific program committee of the Milan conference, we could be biased in commenting on papers dealing with the conference topics. However, it appears useful to add some thoughts here. We should first underline that the consensus statement from this conference in itself does not address the problem of TE vs. uncertainty [7]. This is indeed outside the scope of the statement as it concentrates on models to set performance specifications. The TE and measurement uncertainty concepts were discussed by some of the speakers at the conference and taken forward by one of the Task and Finish Groups (TFG) established during the conference [4].

Westgard's opinion paper contains some criticisms that deserve clarifications [2]. Similarly to a previously published paper on the same topic [8], the paper is mainly a defense of the TE concept the same author promoted 40 years ago more than an objective comment on the need to improve the definition and the quality of performance specifications, which was the main topic of the Milan conference.

We note some misconceptions in the interpretation of the Milan consensus [7], for example, when the author considers the "Milan simplification from 5 to 3 levels of models a *modest adjustment* of the Stockholm guidance". The three models proposed in the consensus document from Milan are not necessarily a hierarchy. The recommendation is to use model 1 or 2 depending on the nature and characteristics of the measurands and depending on their clinical application (i.e. diagnosis, monitoring, etc.). Accordingly, Westgard is ignoring the need, which explains the creation of a TFG by EFLM [4], to allocate different tests or clinical applications of the same test to different models for performance specifications recognized in the consensus statement of the Milan conference, preferentially 1 or 2.

Another ignored aspect is the quality of the information, which in our view is one of the novelties of the Milan conference. In Westgard's paper, biological variability is strongly promoted as a source of analytical performance specifications. This can be a good and usable model for many measurands, but it must be underlined that there is a clear need for evidence-based data and that studies and papers with information about biological variation have to be critically appraised [9, 10]. The utility to elaborate specifications at different levels of quality (i.e. minimum, desirable and optimum) to move, in case, from desirable to minimum quality goals and, in the meantime, ask IVD companies to work for improving the quality of assay performance is also an important aspect [11].

Finally, a confounding approach in the paper of Westgard [2] is the tendency to mix statements contained in the individual papers published in the conference proceedings with agreed concepts that became part of the consensus statement. An example is represented by the content of Petersen's paper [12], which is considered "tout court" the "preference" of the Milan audience. It should be underlined that each of the published papers from the Milan conference represents the opinion of their authors and that there is only one "consensus" paper from the conference [7].

Referring to the TE utility, in his paper Westgard tries to create a sort of "black and white" situation between TE and measurement uncertainty. If we agree that TE may still be useful, for example, in evaluating single EQAS results, there is also a need that all stakeholders working in the field of Laboratory Medicine look carefully to the measurement uncertainty as this help very much in the identification of important sources of

bias [6, 13, 14]. Saying that medical laboratories should take care of TE and IVD manufacturers of measurement uncertainty is a simplification that includes the risk to separate responsibilities that should be conversely strongly integrated [6].

In their contribution, Oosterhuis and Theodorsson further discuss Westgard's opinion by recalling criticisms to the conventional model for deriving allowable TE to be used in assessing quality of laboratory measurements [3]. Oosterhuis previously proposed an alternative model in which the maximum allowable bias and imprecision are interrelated and described in a curve, and the allowable TE calculated from each point of the graph [15]. Using this approach, an overestimation of the allowable TE obtained according to the classical model adding both maximum allowable bias and imprecision was demonstrated. A specific TFG was initiated by EFLM after the Milan conference, commissioned with exploring, developing and coming up with a proposal for how to correctly use the TE concept and how to possibly combine performance specifications for bias and imprecision in a more scientifically sound way [4]. As for other TFGs, this TFG is expected to complete its deliverables within 2 years and clarify when the use of TE can still be useful (e.g. interpretation of single results in EQAS when bias and imprecision effects cannot be separated) or when it should be replaced by the uncertainty estimate. In doing this, it should be remembered that in the world of traceability, where the reference should be the "unbiased" value, the deviation of a laboratory measurement will define its uncertainty [1]. In this scenario, efforts to upgrade the TE framework for defining the allowable uncertainty of a clinical measurement are greatly appreciated.

**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. 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.
2. Westgard JO. Useful measures and models for analytical quality management in medical laboratories. Clin Chem Lab Med 2016;54:223–33.
3. Oosterhuis WP, Theodorsson E. Total error vs. measurement uncertainty: revolution or evolution? Clin Chem Lab Med 2016;54:235–9.
4. Panteghini M, Sandberg S. Defining analytical performance specifications 15 years after the Stockholm conference. Clin Chem Lab Med 2015;53:829–32.
5. Fraser CG. The 1999 Stockholm Consensus Conference on quality specifications in laboratory medicine. Clin Chem Lab Med 2015;53:837–40.
6. Braga F, Panteghini M. Verification of in vitro medical diagnostics (IVD) metrological traceability: responsibilities and strategies. Clin Chim Acta 2014;432:55–61.
7. 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.
8. Westgard JO. Managing quality vs measuring uncertainty. Clin Chem Lab Med 2010;48:31–40.
9. Carobene A. Reliability of biological variation data available in an online database: need for improvement. Clin Chem Lab Med 2015;53:871–7.
10. Bartlett WA, Braga F, Carobene A, Coşkun A, Prusa R, Fernandez-Calle P, et al. A checklist for critical appraisal of studies of biological variation. Clin Chem Lab Med 2015;53:879–85.
11. 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.
12. Petersen PH. Performance criteria based on true and false classification and clinical outcomes. Influence of analytical performance on diagnostic outcome using a single clinical component. Clin Chem Lab Med 2015;53:849–55.
13. 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.
14. Pasqualetti S, Infusino I, Carnevale A, Szőke 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 2015;450:125–6.
15. Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. Clin Chem 2011;57:1334–6.

**Corresponding author: Mauro Panteghini,** Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, via G.B. Grassi 74, Milan 20157, Italy,
E-mail: mauro.panteghini@unimi.it
**Sverre Sandberg:** Norwegian Quality Improvement of Primary Care Laboratories (Noklus), Institute of Global Public Health and Primary Health Care, University of Bergen and Laboratory of Clinical, Biochemistry, Bergen, Norway