

## Opinion Paper

Mauro Panteghini\*

# What the Milan conference has taught us about analytical performance specification model definition and measurand allocation

<https://doi.org/10.1515/cclm-2023-1257>

Received November 7, 2023; accepted January 8, 2024;

published online January 25, 2024

**Abstract:** Analytical performance specifications (APS) represent the criteria that specify the quality required for laboratory test information to satisfy clinical needs. In 2014 the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) considered timely to update the topic of APS by organizing a conference in Milan in which some strategic concepts were proposed. Here I summarize the essential points representing the EFLM Strategic Conference heritage and discuss the approaches that will permit us to become more concrete, including roles and main actions expected from each of involved stakeholders for contributing a quantum leap forward in the way of practicality of Milan consensus about APS.

**Keywords:** analytical performance specifications; standardization; outcome; biological variation; state of the art

## Introduction

In general terms, analytical performance specifications (APS) are the criteria that specify (in numerical terms) the quality required for laboratory test information to satisfy clinical needs. If APS are not objectively defined and fulfilled, the variation in laboratory results may overwhelm the clinical information supplied, even potentially causing a negative effect on patient's outcome. Although 60 years are already passed since the "Tonks' rules" were established [1], the first 35 years of the APS story were spent in a vast and interesting discussion on both sides of the Atlantic Ocean, including some important proposals to set APS that however

did not reach a global consensus [2–6]. Only in 1999, the International Union of Pure and Applied Chemistry (IUPAC)-IFCC conference held in Stockholm was a landmark in trying to achieve a consensus on how APS should be set, with a hierarchy of models being established [7, 8]. In the 10 years following the Stockholm conference there was however a clear realization that the hierarchy of strategies to set APS was not finality as stated in a review of progress made by two of the organizers [9]. It was also considered that the proposed approaches could be modified in the future if significant new ideas were developed and used in practice [10]. Roughly in the same period, the definition and use of the reference system concept for standardization of measurements in Laboratory Medicine became to be closely associated with the setting of targets for measurement uncertainty (MU) to make laboratory results clinically acceptable and establish what degree of quality is needed and what MU can be tolerated without jeopardizing patient safety [11–13]. A debate between the total error (TE) advocates and those conversely supporting the importance of MU, which requires special attention to the bias component and its correction, was also started, including criticisms to the common model employed to derive a limit for the allowable TE using a mathematically incorrect method relying on the sum of mutual exclusive terms [11, 14]. So that, in 2014 the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) considered timely to update the topic of APS by organizing a conference in which some strategic actions and measures should be taken [15].

## The Milan Strategic Conference and its heritage

In November 2014, 215 participants from 41 countries attended the conference, organized by the EFLM in cooperation with the Institute for Reference Materials and Measurements of the European Commission-Joint Research Centre and the Research Centre for Metrological Traceability in Laboratory Medicine (CIRME) of the University of Milan.

---

\*Corresponding author: Mauro Panteghini, Department of Laboratory Medicine, Collegium Medicum, Nicolaus Copernicus University, 9 Skłodowskiej-Curie Street, 85-094 Bydgoszcz, Torun, Poland, Phone: +48 52 585 44 90, E-mail: mauro.panteghini@cm.umk.pl, <https://orcid.org/0000-0002-7147-3433>

All contributions given during the conference were published in June 2015 in a special issue of *Clinical Chemistry and Laboratory Medicine*. In particular, the Consensus Statement describing models for establishing APS has been cited more than 470 times to date (according to Google scholar – <http://scholar.google.com/>) and wide acceptance of the agreed principles has occurred [16–18]. Immediately after the conference, five Task Groups (TG) were established dealing with the main topics of the conference. Among them, one worked on the criteria for assigning different measurands to each of the three Milan models, contributing a fundamental step towards understanding model philosophy and proposing practical principles for how to allocate measurands to different models [19]. Here, it seems useful to employ definitions provided by this TG to discuss about the Strategic Conference heritage.

The outcome-based APS model should be applied “when the measurand has a central and well-defined role in the decision making of a specific disease or a given clinical situation and test results should be interpreted through established decision limits” [19]. Regarding this model, although the Strategic Conference identified in principle two possible types of studies (direct and indirect outcome studies), in the following experience it has become clear that direct outcome studies are unfeasible or impractical, and even not necessary when: a) the clinical decisions associated with the test results are well defined; b) evidence about the diagnostic accuracy of the test to classify patients for these clinical decisions is available and is generalizable to the patient population; c) the consequences of correct/incorrect classification (i.e., true and false positives and negatives) are established [20]. This awareness led to a re-evaluation of studies using indirect methods for assessing the impact of measurement variability on clinical outcome. In doing this, Smith and coworkers showed that a significant number of studies was already available before the Milan conference and should simply be “taken out of the drawer” [21]. Among others, the pioneering contributions of Per Hyltoft Petersen approaching the APS issue by applying distributions of both diseased/non-diseased patient data and simulating the influence of analytical variability on the disease misclassification, still have an important part [22–25]. More recently, other measurands with a central diagnostic role (e.g., cardiac troponin and urine albumin) have been the object of simulation studies evaluating the effects of varying the analytical performance on the patient misclassification [26, 27]. Simulations based on the current thresholds employed in clinical decision making for the definition of the acceptable measurement error have also been done [28].

According to Ceriotti et al. [19], the biological variation (BV)-based APS model should be applied to measurands with

high homeostatic control (e.g., plasma ions) or in a “steady state” status when a subject is in good health (e.g., serum creatinine). Setting APS from BV data uses concepts proposed 25 years ago by Fraser and coworkers [29]. Before and during the Milan conference, criticisms were however raised about the general quality of available BV studies and the need for a set of standard attributes highlighted [30]. Major caveats were related to: a) differences between biology vs. pathophysiology, i.e., the use of diseased individuals for BV data derivation in which the inherent BV may be significantly amplified by the disease (and treatment) related variability [31]; b) statistical drawbacks related to an ignored non-Gaussian distribution of the data (needing mathematical transformations for calculations), resulting in too high biological CV estimates [32] or to a poor study power to deliver robust data with narrow confidence intervals [33]; c) analytical issues, such as different method selectivity or insufficient method sensitivity to detect measurands present in plasma at very low concentrations *in all samples of all subjects* enrolled in the BV protocols [27, 34]. This is why a dedicated TG was established by EFLM for operating a careful appraisal of available BV studies by using a detailed checklist and generate an exhaustive database on the EFLM website [35–38]. It has been however noted that it is not acceptable to use the BV-based model to derive APS for all measurands just because the BV information is now more easily obtainable from the mentioned database: in particular, the model 2 should not be used for measurands having not sufficient steady state status, such as, for instance, most hormones [39].

Defined by the Milan consensus as the highest level of analytical performance technically achievable, the “state of the art” (SA) should be used when a measurand has neither central diagnostic role nor strict homeostatic control [19]. For example, this model does apply for most urinary measurands, for which the concentrations vary to maintain the corresponding plasmatic concentrations stable, compensating for dietary provision, water supplementation or loss, etc. In addition, the model 3 can be temporarily used also for those measurands still waiting for the definition of outcome-based APS or while waiting for robust BV data [19]. On the other hand, it is important to pointing up that the myth of SA as a ‘rescue’ model when APS correctly obtained with other more appropriate models for a certain measurand appear too stringent should be on principle dismantled [40]. Even if this opinion is not shared by everyone, using APS derived from the correct allocation of measurands in different models has been shown helpful in identifying measurands that need analytical improvement for their clinical use. The case of plasma electrolyte measurements recently described showed the efficacy of using objectively derived APS in

driving laboratories to improve the quality of provided results [41]. Recognized drawbacks for this model, partially already highlighted during the EFLM conference, are: a) the difficulty to define how good the 'highest' is; b) some dependency on industry defined quality, and, more importantly, c) the lack of direct relationship between what is analytically achievable and what is clinically needed, potentially leading to unnecessary improvements. A proposal to define SA best MU performance according to the definition provided by the Milan conference and the ISO/TS 20914 technical guidance for the estimation of MU [42], taking serum C-reactive protein as an illustrative measurand, has been published [40, 43].

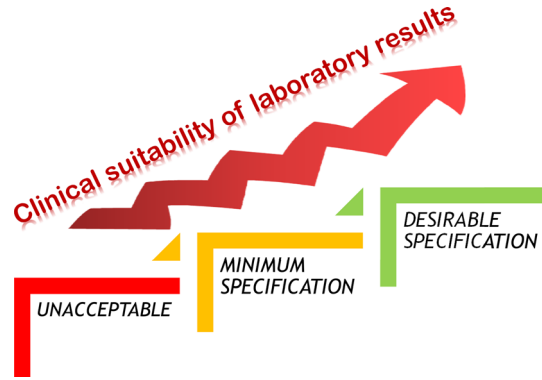
Table 1 summarizes the essential points representing the EFLM Strategic Conference heritage. One can consider as a part of this heritage also the fact that, in the last few years, International Organization for Standardization (ISO) standards and guidance dealing with metrological traceability in Laboratory Medicine started to explicitly emphasize the importance of APS. While ISO 17511:2020 made a direct reference to the Milan conference in the Note 1 of clause 4.3.2 [44], ISO/TS 20914:2019 widely summarized the conference deliberations under the paragraph 5.2 dealing with the concept of maximum allowable MU [42].

## Moving from theoretical principles to practice

Nine years after the Milan conference it is now time to identify the approaches that permit us to become more concrete. First, the concepts described in the previous paragraph should be applied for all measurands used in the clinical setting. We should extend the list: any single

**Table 1:** Essential items as part of the EFLM Strategic Conference heritage.

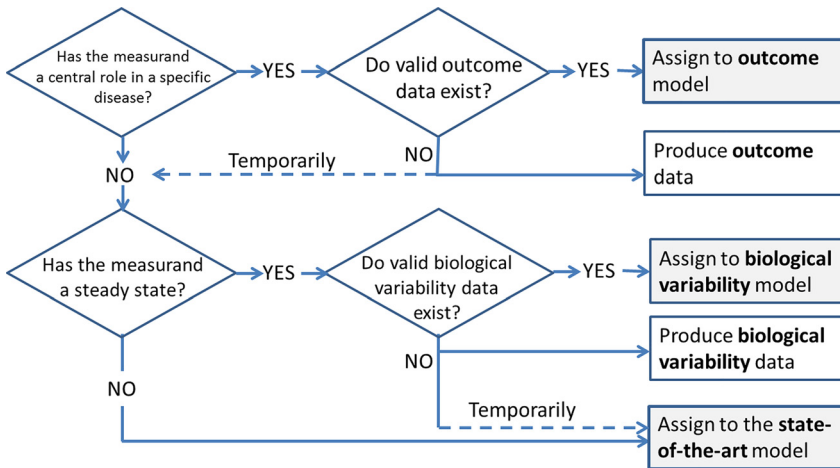
- 
- Three models to set analytical performance specifications (APS): outcome, biology, state of the art.
  - The models use different principles and do not necessarily constitute a hierarchy.
  - Some models are better suited for certain measurands than for others, and the attention should primarily direct toward the measurand and its biological and clinical characteristics.
  - APS can be different for different applications of the same test, but if a test is used for multiple purposes the strictest APS should take precedence.
  - Newly defined APS should derive from high quality studies and updated data.
- 



**Figure 1:** The stairway of analytical performance specifications: grading quality levels helps to define priorities in ameliorating actions by stakeholders sharing the responsibility of laboratory performance, i.e., *in vitro* diagnostic manufacturers and medical laboratories themselves.

laboratory measurand should be allocated to a specific APS model. Second, based on the chosen model, APS values should be established and put into practice. This is an ongoing process that would require some kind of official endorsement from professional laboratory bodies. In doing this, we should acknowledge that not all the established desirable APS will be immediately reachable, but this approach will highlight which limitations of the current technology should be prioritized and solved. Very helpful is grading different quality levels of APS to define priorities in ameliorating actions (Figure 1) [17]. Even if the derivation approach of this grading has sometimes little scientific basis, the possibility to elaborate APS at different levels of quality to move, in case, from desirable to minimum APS when the requirements are particularly stringent may help to accept the described theory and, on the other hand, represent a stimulus for involved stakeholders to work for improving the quality of assay performance when needed.

Figure 2 depicts the workflow for assignment of a measurand to an APS model as defined by Ceriotti et al. following the Milan conference [19]. The recently published APERTURE study, a project for establishing Analytical Performance Specifications for Measurement Uncertainty of most common laboratory measurands, has used this flowchart to allocate different measurands to one of the three APS models [18, 27, 39, 45–48]. Afterwards, both minimum and desirable quality levels of APS for standard MU of clinical samples were defined by using information obtained from available literature preliminarily checked in terms of robustness. A list of information related to the selected measurands, model allocation, and APS definition is available in Table 4 of reference [18].



**Figure 2:** Workflow for assignment of a measurand to one of the analytical performance specification models as defined in the 2014 Milan conference. Adapted from Ceriotti et al. [19].

The related figures were then adopted by the Task Force on Reference Measurement System Implementation of the Joint Committee on Traceability in Laboratory Medicine (JCTLM) in their ongoing work to produce a synopsis of JCTLM-listed higher-order certified reference materials and reference measurement procedures for selected measurands, including their main characteristics for implementing metrological traceability and fulfilling (or not) the APS for suitable MU [49].

## What further developments are expected in the APS field?

Drugs represent a special category of measurands that probably needs a dedicated approach for deriving APS, based on fundamental pharmacokinetic theory and average elimination half-life of the drug [50, 51]. Based on these concepts, Braga et al. have proposed a kind of hybrid model, mixing some characteristics of Milan models 1 and 2 [45]. Indeed, although concentrations of drugs do not fluctuate randomly around a homeostatic set point, the proposed approach has a relationship with biological knowledge. On the other hand, therapeutic drug monitoring (TDM) is also linked to the patient outcome in defining the concentrations of drug which are potentially toxic or when the treatment can be ineffective. Using this conceptual framework, APS for MU in TDM of immunosuppressive drugs were recently established [52].

In the next future, computerised approaches using the framework of simulation studies will become more popular and definition of APS according to indirect outcome-based studies less difficult [21]. As an example, by using a modification of the typical error model published in 2001 by Boyd and Bruns [53], the EFLM Working Group on

Accreditation and ISO/CEN Standards has quite recently made up a tool for setting APS for MU of measurands that should be allocated to Milan model 1 [54]. This process can be directly implemented by laboratory professionals using a freely available web application (<https://hikmetc-apscalculator.streamlit.app/>), providing that users have test value data of the population of clinical interest and the evidence that the underlying data are a reliable proxy for the truth [21]. Multiple verifications of the software by other investigators are expected in the next future. Other studies have employed contour plots to present findings from which to derive APS information according to a given rate of clinical misclassification [55, 56].

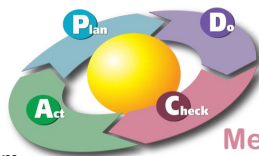
An important aspect that relates with practical application of objectively defined APS is their introduction as mandatory feature in the External Quality Assessment (EQA) programs for the evaluation of the performance of participating laboratories in terms of inter-assay harmonization of the provided results and, ultimately, of the suitability of laboratory measurements in clinical settings [11, 13, 18, 34, 57–60]. Just after the Milan conference, Jones et al. highlighted the wide variation in the definition of APS used by EQA programs and the huge difference in the impact of provided information [61]. Consensus has not yet been reached on this topic. However, taking as reference the almost classic EQA program categorization proposed by Miller and co-workers [62], it is probably time to integrate the EQA category 1–2 by adding a letter that identifies which model is employed by the EQA organizers to define APS, i.e., 1–2 A when one of the Milan models is used; and 1–2 B when other models, e.g., statistic-based, are employed [58, 59]. Validation/acceptability criteria for the Internal Quality Control should also be redesigned by using appropriately derived APS [18, 63].

Endorse and adopt APS defined according to the Milan principles, and complete the allocation of laboratory measurands to appropriate APS models

### Laboratory profession

### All stakeholders

Work together for optimum patient care by applying suitable APS



Give more focus to the minority of tests that need urgent improvement and operate to reduce as much as possible the measurement uncertainty of calibrators, especially when APS are stringent

### IVD manufacturers

### Medical laboratories

Supervise the IVD-MD quality and their own performance using appropriate IQC and EQA approaches, which should include objectively derived APS

**Figure 3:** Roles and main actions expected from each stakeholder for contributing a quantum leap forward in the way of practicality of Milan consensus about analytical performance specifications; APS, analytical performance specifications; IVD-MD, *in vitro* diagnostic-medical device; IQC, internal quality control; EQA, external quality assessment.

## Conclusions

As discussed in this paper, in the last nine years a considerable work to underpin APS according to the Milan consensus has been performed. However, we should generally recognize that many aspects remain predominantly theoretical. It is still a question whether APS theory as agreed during the Milan conference will be enough to convince stakeholders that change is needed. Therefore, now we need to be more functional and make a quantum leap forward in the way of practicality. According to roles and actions described in Figure 3, stakeholders are expected to work together in the field of APS for improving the quality and suitability of contributions of Laboratory Medicine to patient care. Particularly, pressure from the laboratory professionals will be crucial to invoke change and an APS list supported by the professional bodies in the field should be welcomed for helping to identify which tests (a minority) require an actual improvement.

**Research ethics:** Not applicable.

**Informed consent:** Not applicable.

**Author contributions:** The author has accepted responsibility for the entire content of this manuscript and approved its submission.

**Competing interests:** The author states no conflict of interest.

**Research funding:** None declared.

**Data availability:** Not applicable.

## References

1. Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. *Clin Chem* 1963;9: 217–33.
2. Cotlove E, Harris EK, Williams GZ. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. III. Physiological and medical implications. *Clin Chem* 1970; 16:1028–32.
3. Elevitch FR, editor. College of American Pathologists Conference II (1976): Analytical goals in clinical chemistry. Skokie, IL: College of American Pathologists; 1977.
4. Fraser CG, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems in clinical chemistry. *Eur J Clin Chem Clin Biochem* 1992;30:311–7.
5. Stockl D, Baadenhuisjen H, Fraser CG, Libeer JC, Hyltoft Petersen P, Ricos C. Desirable routine analytical goals for quantities assayed in serum. *Eur J Clin Chem Clin Biochem* 1995;33:157–69.
6. Hyltoft Petersen P, Ricós C, Stöckl D, Libeer J-C, Baadenhuisjen H, Fraser CG, et al. Proposed guidelines for the internal quality control of analytical results in the medical laboratory. Discussion paper from the members of the external quality assessment (EQA) working group A on the analytical quality goals in laboratory medicine. *Eur J Clin Chem Clin Biochem* 1996;34:983–99.
7. Kallner A, McQueen M, Heuck C. The Stockholm consensus conference on quality specifications in laboratory medicine, 25–26 April 1999. *Scand J Clin Lab Invest* 1999;59:475–585.
8. Fraser CG. The 1999 Stockholm Consensus Conference on quality specifications in laboratory medicine. *Clin Chem Lab Med* 2015;53: 837–40.
9. Hyltoft Petersen P, Fraser CG. Strategies to set global analytical quality specifications in laboratory medicine: 10 years on from the Stockholm Consensus Conference. *Accred Qual Assur* 2010;15:323–30.
10. Hyltoft Petersen P, Sandberg S, Fraser CG. Do new concepts for deriving permissible limits for analytical imprecision and bias have any advantages over existing consensus? *Clin Chem Lab Med* 2011;49: 637–40.
11. 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.
12. Panteghini M. Implementation of standardization in clinical practice: not always an easy task. *Clin Chem Lab Med* 2012;50:1237–41.
13. Braga F, Panteghini M. Verification of *in vitro* medical diagnostics (IVD) metrological traceability: responsibilities and strategies. *Clin Chim Acta* 2014;432:55–61.
14. Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. *Clin Chem* 2011;57:1334–6.

15. Panteghini M, Sandberg S. Defining analytical performance specifications 15 years after the Stockholm conference. *Clin Chem Lab Med* 2015;53:829–32.
16. 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.
17. Panteghini M, Ceriotti F, Jones G, Oosterhuis W, Plebani M, Sandberg S, Task Force on Performance Specifications in Laboratory Medicine of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). Strategies to define performance specifications in laboratory medicine: 3 years on from the Milan Strategic Conference. *Clin Chem Lab Med* 2017;55:1849–56.
18. Panteghini M. Redesigning the surveillance of in vitro diagnostic medical devices and of medical laboratory performance by quality control in the traceability era. *Clin Chem Lab Med* 2023;61:759–68.
19. Ceriotti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, et al., EFLM Task and Finish Group on Allocation of laboratory tests to different models for performance specifications (TFG-DM). 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.
20. Horvath AR, Bossuyt PM, Sandberg S, St John A, Monaghan PJ, Verhagen-Kamerbeek WD, et al. Setting analytical performance specifications based on outcome studies – is it possible? *Clin Chem Lab Med* 2015;53:841–8.
21. Smith AF, Shinkins B, Hall PS, Hulme CT, Messenger MP. Toward a framework for outcome-based analytical performance specifications: a methodology review of indirect methods for evaluating the impact of measurement uncertainty on clinical outcomes. *Clin Chem* 2019;65:1363–74.
22. Wiggers P, Dalhøj J, Hyltoft Petersen P, Blaabjerg O, Hørder M. Screening for haemochromatosis: influence of analytical imprecision, diagnostic limit and prevalence on test validity. *Scand J Clin Lab Invest* 1991;51:143–8.
23. Hyltoft Petersen P, Brandslund I, Jørgensen L, Stahl M, de Fine Olivarius N, Borch-Johnsen K. Evaluation of systematic and random factors in measurements of fasting plasma glucose as the basis for analytical quality specifications in the diagnosis of diabetes. 3. Impact of the new WHO and ADA recommendations on diagnosis of diabetes mellitus. *Scand J Clin Lab Invest* 2001;61:191–204.
24. Hyltoft Petersen P, Klee GG. Influence of analytical bias and imprecision on the number of false positive results using guideline-driven medical decision limits. *Clin Chim Acta* 2014;430:1–8.
25. Nielsen AA, Petersen PH, Green A, Christensen C, Christensen H, Brandslund I. Changing from glucose to HbA1c for diabetes diagnosis: predictive values of one test and importance of analytical bias and imprecision. *Clin Chem Lab Med* 2014;52:1069–77.
26. Ko DH, Lee SW, Hyun J, Kim HS, Park MJ, Shin DH. Proposed imprecision quality goals for urinary albumin/creatinine ratio. *Ann Lab Med* 2018;38:420–4.
27. Krintus M, Panteghini M. Judging the clinical suitability of analytical performance of cardiac troponin assays. *Clin Chem Lab Med* 2023;61:801–10.
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. 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.
30. Braga F, Panteghini M. Generation of data on within-subject biological variation in laboratory medicine: an update. *Crit Rev Clin Lab Sci* 2016;53:313–25.
31. Carlsen S, Petersen PH, Skeie S, Skadberg Ø, Sandberg S. Within-subject biological variation of glucose and HbA(1c) in healthy persons and in type 1 diabetes patients. *Clin Chem Lab Med* 2011;49:1501–7.
32. Braga F, Panteghini M. Biologic variability of C-reactive protein: is the available information reliable? *Clin Chim Acta* 2012;413:1179–83.
33. Røraas T, Petersen PH, Sandberg S. Confidence intervals and power calculations for within-person biological variation: effect of analytical imprecision, number of replicates, number of samples, and number of individuals. *Clin Chem* 2012;58:1306–13.
34. Braga F, Panteghini M. Standardization and analytical goals for glycated hemoglobin measurement. *Clin Chem Lab Med* 2013;51:1719–26. Erratum in: *Clin Chem Lab Med* 2014;52:759.
35. Aarsand AK, Røraas T, Fernandez-Calle P, Ricos C, Díaz-Garzón J, Jonker N, et al., European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation and Task and Finish Group for the Biological Variation Database. The biological variation data critical appraisal checklist: a standard for evaluating studies on biological variation. *Clin Chem* 2018;64:501–14.
36. Aarsand AK, Røraas T, Bartlett WA, Coşkun A, Carobene A, Fernandez-Calle P, et al., European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group on Biological Variation. Harmonization initiatives in the generation, reporting and application of biological variation data. *Clin Chem Lab Med* 2018;56:1629–36.
37. Sandberg S, Carobene A, Aarsand AK. Biological variation – eight years after the 1st Strategic Conference of EFLM. *Clin Chem Lab Med* 2022;60:465–8.
38. Sandberg S, Carobene A, Bartlett B, Coskun A, Fernandez-Calle P, Jonker N, et al. Biological variation: recent development and future challenges. *Clin Chem Lab Med* 2023;61:741–50.
39. Braga F, Panteghini M. Performance specifications for measurement uncertainty of common biochemical measurands according to Milan models. *Clin Chem Lab Med* 2021;59:1362.
40. Braga F, Panteghini M. Derivation of performance specifications for uncertainty of serum C-reactive protein measurement according to the Milan model 3 (state of the art). *Clin Chem Lab Med* 2020;58:e263–5.
41. Pasqualetti S, Chibireva M, Borrillo F, Braga F, Panteghini M. Improving measurement uncertainty of plasma electrolytes: a complex but not impossible task. *Clin Chem Lab Med* 2020;59:e129–32.
42. ISO/TS 20914:2019. Medical laboratories – practical guidance for the estimation of measurement uncertainty, 1st ed. Geneva, Switzerland: International Organization for Standardization (ISO); 2019.
43. Borrillo F, Panteghini M. Current performance of C-reactive protein determination and derivation of quality specifications for its measurement uncertainty. *Clin Chem Lab Med* 2023;61:1552–7.
44. ISO 17511:2020. In vitro diagnostic medical devices – requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. Geneva, Switzerland: International Organization for Standardization (ISO); 2020.
45. Braga F, Pasqualetti S, Borrillo F, Capoferri A, Chibireva M, Rovegno L, et al. Definition and application of performance specifications for measurement uncertainty of 23 common laboratory tests: linking theory to daily practice. *Clin Chem Lab Med* 2023;61:213–23.
46. Borrillo F, Pasqualetti S, Panteghini M. Measurement uncertainty of thyroid function tests on a chemiluminescent microparticle immunoassay system needs to be improved. *J Appl Lab Med* 2023;8:420–2.

47. Rovegno L, Civera E, Infusino I, Panteghini M. State of the art of measurement uncertainty of serum ferritin. *Clin Chem Lab Med* 2024;62:e6–8.
48. Borrillo F, Capoferri A, Rovegno L, Panteghini M. Closed analyser lids do not reduce the measurement uncertainty of serum total carbon dioxide. *Ann Clin Biochem* 2023 Nov 18: 45632231216598. <https://doi.org/10.1177/00045632231216598> [Epub ahead of print].
49. Panteghini M, Braga F, Camara JE, Delatour V, Van Uytvanghe K, Vesper HW, et al., JCTLM Task Force on Reference Measurement System Implementation. Optimizing available tools for achieving result standardization: value added by Joint Committee on Traceability in Laboratory Medicine (JCTLM). *Clin Chem* 2021;67: 1590–605.
50. Fraser CG. Desirable standards of performance for therapeutic drug monitoring. *Clin Chem* 1987;33:387–9.
51. Stewart MJ, Fraser CG. Desirable performance standards for assays of drugs. *Ann Clin Biochem* 1989;26:220–6.
52. Cattaneo D, Panteghini M. Analytical performance specifications for measurement uncertainty in therapeutic monitoring of immunosuppressive drugs. *Clin Chem Lab Med* 2024;62:e81–3.
53. Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. *Clin Chem* 2001;47:209–14.
54. Çubukçu HC, Vanstapel F, Thelen M, van Schrojenstein Lantman M, Bernabeu-Andreu FA, Meško Brguljan P, et al. APS calculator: a data-driven tool for setting outcome-based analytical performance specifications for measurement uncertainty using specific clinical requirements and population data. *Clin Chem Lab Med* 2024;62: 597–607.
55. Lyon AW, Kavsak PA, Lyon OA, Worster A, Lyon ME. Simulation models of misclassification error for single thresholds of high sensitivity cardiac troponin I due to assay bias and imprecision. *Clin Chem* 2017;63:585–92.
56. 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.
57. Ferraro S, Braga F, Panteghini M. Laboratory medicine in the new healthcare environment. *Clin Chem Lab Med* 2016;54:523–33.
58. Infusino I, Frusciante E, Braga F, Panteghini M. Progress and impact of enzyme measurement standardization. *Clin Chem Lab Med* 2017;55: 334–40.
59. 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.
60. Badrick T, Jones G, Miller WG, Panteghini M, Quintenz A, Sandberg S, et al. Differences between educational and regulatory external quality assurance/proficiency testing schemes. *Clin Chem* 2022;68:1238–44.
61. Jones GRD, Albarede S, Kessler D, MacKenzie F, Mammen J, Pedersen M, et al. Analytical performance specifications for external quality assessment – definitions and descriptions. *Clin Chem Lab Med* 2017;55:949–55.
62. Miller WG, Jones GR, Horowitz GL, Weykamp C. Proficiency testing/ external quality assessment: current challenges and future directions. *Clin Chem* 2011;57:1670–80.
63. Braga F, Pasqualetti S, Aloisio E, Panteghini M. The internal quality control in the traceability era. *Clin Chem Lab Med* 2021;59:291–300.