Opinion Paper

Sverre Sandberg*, Abdurrahman Coskun, Anna Carobene, Pilar Fernandez-Calle, Jorge Diaz-Garzon, William A. Bartlett, Niels Jonker, Kornelia Galior, Elisabet Gonzales-Lao, Isabel Moreno-Parro, Berta Sufrate-Vergara, Craig Webster and Aasne K. Aarsand, on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Task Group for the Biological Variation Database

Analytical performance specifications based on biological variation data – considerations, strengths and limitations

https://doi.org/10.1515/cclm-2024-0108
Received January 22, 2024; accepted March 4, 2024;
published online March 20, 2024

Abstract: Analytical performance specifications (APS) are typically established through one of three models: (i) outcome studies, (ii) biological variation (BV), or (iii) state-of-the-art. Presently, The APS can, for most measurands that have a stable concentration, be based on BV. BV based APS, defined for imprecision, bias, total allowable error and allowable measurement uncertainty, are applied to many different processes in the laboratory. When calculating APS, it is important to consider the different APS formulae, for what setting they are to be applied and if they are suitable for the intended purpose. In this opinion paper, we elucidate the background, limitations, strengths, and potential intended applications of the different BV based APS formulas. When using BV data to set APS, it is important to consider that all formulae are contingent on accurate and relevant BV estimates. During the last decade, efficient procedures have been established to obtain reliable BV estimates that are presented in the EFLM biological variation database. The database publishes detailed BV data for numerous measurands, global BV estimates derived from meta-analysis of quality-assured studies of similar study design and automatic calculation of BV based APS.

Keywords: analytical performance specifications; biological variation; precision; bias; total error; measurement uncertainty

***Corresponding author: Sverre Sandberg**, Norwegian Organization for Quality Improvement of Laboratory Examinations (Noklus), Haraldsplass Deaconess Hospital, Bergen, Norway; Department of Medical Biochemistry and Pharmacology, The Norwegian Porphyria Centre, Haukeland University Hospital, Bergen, Norway; and Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway,
E-mail: sverre.sandberg@noklus.no
Abdurrahman Coskun, Department of Medical Biochemistry Atasehir, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Istanbul, Türkiye. https://orcid.org/0000-0002-1273-0604
Anna Carobene, Laboratory Medicine, IRCCS San Raffaele Scientific Institute, Milan, Italy
Pilar Fernandez-Calle, **Jorge Diaz-Garzon**, **Isabel Moreno-Parro and Berta Sufrate-Vergara,** Laboratory Medicine Department, La Paz University Hospital, Madrid, Spain. https://orcid.org/0000-0002-3171-1505 (J. Diaz-Garzon)
William A. Bartlett, Biomedical Engineering, School of Engineering and Science, University of Dundee, Dundee, Scotland. https://orcid.org/0000-0002-4105-7471
Niels Jonker, Certe, Wilhelmina Ziekenhuis Assen, Assen, The Netherlands
Kornelia Galior, Department of Pathology and Laboratory Medicine, Emory University, Atlanta, GA, USA
Elisabet Gonzales-Lao, Quality and Patient Safety Department, Consorci Sanitari de Terrassa University Hospital, Barcelona, Spain
Craig Webster, Department of Biochemistry, Immunology and Toxicology, University Hospitals Birmingham, Birmingham, UK
Aasne K. Aarsand, Norwegian Organization for Quality Improvement of Laboratory Examinations (Noklus), Haraldsplass Deaconess Hospital, Bergen, Norway; and Department of Medical Biochemistry and Pharmacology, The Norwegian Porphyria Centre, Haukeland University Hospital, Bergen, Norway

Introduction

Biological variation (BV) refers to the variation that is observed in the concentration or activity of a measurand within an individual or between individuals. It is an important concept in laboratory medicine for many reasons, one of them is that analytical performance specifications (APS) can, for most measurands that have a stable concentration, be based on BV. BV data are also used in the interpretation of laboratory test results, such as for defining reference change values, personalized reference intervals, the index of individuality and other applications. There are various concepts enumerating BV data; these include (i) the within-subject BV (CV_I), which is defined as the variation of

Open Access. © 2024 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

the concentration/activity of a measurand around a homeostatic set point within a single individual in steady state, estimated in a group of individuals, and (ii) the between-subject BV (CV_G), which refers to the variation between the homeostatic set points of different individuals. APSs are defined for many different properties in the laboratory and are commonly calculated for imprecision, bias, total allowable error (TEa) and allowable measurement uncertainty (MAu). When setting APS these can be aspirational or more pragmatic, for use in "daily life".

However, importantly, when using BV data for setting APS and other laboratory applications, this requires BV data to be of high quality and fit for purpose [1]. There have been numerous BV studies published, but the quality of these studies, both in terms of execution and reporting, vary. In the last decade, various initiatives have been undertaken to improve on the quality of available BV data. One major initiative is the establishment of the EFLM biological variation database [2], which provides detailed and quality-appraised BV data for numerous measurands, and automatic calculation of APS, based on global BV estimates derived from meta-analysis of quality-assessed studies. However, it is essential that users of the database understand how to best select appropriate APS for use in their laboratory and the limitations in available BV data. In this opinion article, we examine which APS definitions are based on BV data, describe how to establish BV-based APS and discuss the strengths and limitations of the different approaches included in the EFLM biological variation database. We also provide an overview of the establishment of the EFLM BV database and its current state.

Analytical performance specifications based on biological variation

There are many different strategies available to set APS, which have evolved over time. The main criterium that APS should be based on BV was first formally accepted in 1976 in a College of American Pathologists conference [3]. It dealt with how to set APS for imprecision, describing two different scenarios:

1) "Group screening where an individual is to be selected from a population, a specification for imprecision (CV_A)" was defined as:

$$CV_A = 0.5 \times \sqrt{CV_I^2 + CV_G^2}$$

and 2) "For individual single and multipoint testing, in which an individual is evaluated on the basis of discrimination values":

$$CV_A = 0.5 \times CV_I.$$

In 1988 Gowans proposed the formula for acceptable bias as $0.25 \times (CV_I^2 + CV_G^2)^{1/2}$, defined for using harmonised/transferable reference intervals between laboratories, presupposing imprecision of "almost zero" [4]. In a 1993 publication [5], Fraser and Petersen proposed the formulae that are most commonly used today for acceptable imprecision, bias and TEa, where the latter was intended to be used for External Quality Assessment (EQA). In a review paper about APS published the same year, Petersen et al. again emphasised that if imprecision is not zero, the factor 0.25 in the bias formula should be modified [6]. In the same paper, it was underlined that "The basic principle (for using BV estimates to set APS) is that the analytical errors should have a minimal influence compared to: (i) the biological within-subject variation, and (ii) the total biological (combined within- and between-subject) variation". In 1999, the Stockholm Conference on "Strategies to set global analytical quality specifications in laboratory medicine" [7] advocated the ubiquitous application of a hierarchical structure of approaches, based on a model proposed by Fraser and Petersen [8]. In one of the levels, level 2b, it was stated that "data based on components of biological variation" should be used to set APS. In the two editorials from 1991 to 1999 by Fraser and Petersen [8, 9] it was further elaborated that APS for imprecision could be determined as CV_A=0.5 × CV_I, and APS for bias could be determined as $0.25 \times (CV_I^2 + CV_G^2)^{1/2}$. Furthermore, it was stated that imprecision and bias should be judged separately.

In 2014, in the 1st Strategic Conference of the EFLM, the concept of three different models to set APS was agreed upon [10]. These included (i) model 1; based on the effect of analytical performance on clinical outcomes with sub-model 1a using direct outcome studies and sub-model 1b using outcome studies based on modelling, (ii) model 2; based on components of BV and (iii) model 3; based on state of the art. When the definition of APS was readdressed in the Strategic Conference, it was explicitly stated that the main reason to set APS based on BV was that "analytical noise should be small compared with biological variation" [10, 11]. A proposal for how to allocate different measurands to different models, including the BV model, has been published [12]. Presently, due to the lack of data to establish APS based on model 1, APS can, for most measurands that have a stable concentration, be based on BV. In the following, the background and rationale for the different BV-based APS presented in the EFLM database on BV [2] are summarized and discussed. There are also additional APS formulae utilizing BV data that will not be addressed here.

APS for analytical imprecision (CV_A)

As described above, APS for analytical imprecision is, from a historical point of view, the "oldest" BV-based APS approach [3, 6, 13]. In this, the APS for CV_A is simply set as equal to $0.5 \times CV_I$, with the rationale that the noise (analytical imprecision) then will be small (12 %) compared to the CV_I [14]. The set percentage for acceptable noise is, of course, arbitrary and was possibly chosen simply because "0.5" is a "nice and easy" number. However, it is seldom specified in what "precision" situations the formula is applicable for; is it e.g. for repeatability precision, intermediate precision, or reproducibility precision?

APS for bias

Bias is a component of measurement error that in replicate measurements remains constant or varies within a defined scope [15]. Ideally, in laboratory processes, a significant bias should be eliminated, and efforts are directed into minimizing the bias between measurement procedures [9]. However, from a practical point of view, it may not be possible to remove significant biases completely, and it may then be useful to define an APS as to what constant systematic differences can be acceptable, without compromising clinical decisions. The most commonly used APS for bias ($0.25 \times (CV_I^2 + CV_G^2)^{1/2}$) had, as described above, the aim to set "analytical goals required for the successfully transfer of reference intervals between laboratories within a specified limited geographical area, with a population homogeneous for the quantities" [4]. Furthermore, the authors stated that presuppositions for using this formula was that "the maximum acceptable percentage of the population outside the limit for the 0.90 confidence interval of each of the reference limits is 4.6 % for a population sample size of 120" and that "bias should not exceed $0.25 \times (CV_I^2 + CV_G^2)^{1/2}$, when analytical imprecision is almost nothing, or that the percentage of the population outside each reference limit should not exceed 4.6 % for a combination of analytical imprecision and bias". Thus, the APS for bias derived by this formula, as it is now commonly used in many laboratories, is overestimated since imprecision is never zero and rarely negligible compared to any bias. Another limitation of this formula is that it assumes a Gaussian distribution of the data forming the basis for the reference interval, which is seldom the case. In the work defining this APS, it was underlined that the bias formula and the imprecision formula should be used separately. However, this is ignoring the problem that if they are used separately in one laboratory, the total acceptable "error" in the laboratory is actually the combination of the two.

It can be debated if this formula is actually "based on biological variation" since although estimates of within- and between-subject biological variation are included, the rationale behind it is an "expert opinion" i.e. that no more than 4.6 % of the population should be outside each reference limit. Thus, this is actually more of a model 1b approach [10]. However, when the formula is used today, it is often used for setting APS for all kinds of biases, presumably with the rationale to minimise analytical noise relative to both within- and between-subject BV [6]. From this point of view, it is a BV based approach.

APS for total allowable error (TEa) and allowable measurement uncertainty (MAu)

A laboratory worker monitoring e.g. a measuring system may find separate APS for bias and imprecision useful for quality control purposes. Users of laboratory test results, however, will not be likely to care if the uncertainty of a test results is caused by imprecision or bias, as from a clinical point of view, the total error or uncertainty of the result is what matters. The concept of TE, although mentioned by Eisenhart in 1963 [16], was first elaborated by Westgard in 1972 [17]. MU, as presented in GUM in 1993 [18] and TE represent different paradigms in metrology. Kallner aptly summarized the difference as follows: "MU defines an interval around a measured value, where the true value can be found with a certain probability. TE defines an interval around a true value, where the measured value can be found with a certain probability" [19]. TE is based on the error concept where a reference value, serving as an unbiased estimate of the true value, is known. In MU, however, we are talking about the uncertainty of the measurement result, with a value that can be traced back to a standard of a higher order. In VIM3 [15], it is stated that "No rule can be derived on how they (random and systematic error) combine to form the total error of any given measurement result, usually taken as the estimate. Usually, only an upper limit of the absolute value of the total error is estimated, sometimes loosely named "uncertainty"". It is often stated that TE is easier to calculate than MU. This might be true if the "bottom up" approach, which is based on the combined uncertainty where the uncertainty of each individual measurement step across the entire metrological traceability chain is identified, quantified and added, is used for calculating MU. However, if a "top down" approach, which is based on estimates of the

measurement uncertainty from repeated measurements, typically by evaluating internal quality control data is used, MU can be rather easy to calculate [20]. A comparison of the use of error and measurement uncertainty methods in laboratory medicine can be found in refs. [21, 22].

APS for total allowable error (TEa) estimated as a combination of allowable imprecision and allowable bias

The total allowable error (TEa) formula, as shown below, combines both bias and imprecision to define the APS. It was proposed by Fraser CG and Petersen PH, first published in 1993, and it was intended to be used to define acceptable limits for EQA schemes based on single determination of circulated survey material [5]. In EQA, ideally the measurement accuracy of one single result is compared with a target value from a reference measurement procedure. Whether the deviating result is due to poor precision or to poor trueness is not known. The formula is, however, simple and easy to remember.

$$\mathrm{TEa} < 1.65 \times (0.5 \times \mathrm{CV_I}) + 0.25 \times \left(\mathrm{CV_I^2} + \mathrm{CV_G^2}\right)^{1/2}$$

It combines the allowable imprecision with allowable bias, which, as previously described, has no mathematical basis [15, 23], and it results in an overestimation of TEa. Nevertheless, it is widely used, including for general laboratory applications. In a personal communication from Professor Callum G Fraser in relation to this manuscript, he writes: “In my opinion, these concepts (TEa) are outdated. A particular concern is the use of an APS for bias, which is based, in part, on CV_G. While CV_I seems to be able to be regarded as a “constant”, CV_G is dependent on the population studied. Thus, the APS for bias will vary dependent on the population used to derive the components of analytical and biological variation. At this point in time, opinion seems to have generally returned to my view of the 1970s and 1980s [24] that bias ought to be eliminated, and the use of measurement uncertainty has become widespread and is required by accreditation bodies. Importantly, no bias means that common reference intervals can be applied throughout a geographical area if the methods examine the same measurand and allows the correct use of fixed numerical criteria for interpretation. If then, bias should be eliminated, the APS for MU clearly should be based only on CV_I”. In laboratory practice, bias is often not eliminated for many measurands measured using different measurement procedures. Laboratories should critically consider when the TEa approach is applicable to use.

APS for allowable standard measurement uncertainty (MAu) and for allowable expanded measurement uncertainty (MAU)

In the MU paradigm, bias should be corrected or eliminated, and all the remaining sources of variation added linearly as variances. A measurement result therefore can comprise two uncertainties: the uncertainty associated with bias correction and the uncertainty due to imprecision. The APS for allowable standard measurement uncertainty (MAu) can be set as $0.5 \times CV_I$, similar to the APS for imprecision, and the allowable expanded measurement uncertainty (MAU) is $k \times 0.5 \times CV_I$. The “k” is the coverage factor, for example, 2 or 3, to obtain a certain confidence level (approximately 95 or 99 % respectively). The coverage factor most in use is 2, reflecting the “usual” 95 % confidence interval (CI) used in laboratory medicine. In this situation, the APS MAU can be calculated as $2 \times 0.5 \times CV_I$. These formulas are based on BV, with the aim to minimize analytical noise compared to the BV. They can be used to verify that the estimated MU is within the set APS and to facilitate communication of the uncertainty of a patient result to clinicians and patients. It may be more challenging to use for monitoring and controlling instruments in the laboratory, where detecting bias and imprecision separately may be preferable. The MAu can in many cases be considered as an APS for imprecision performed under intermediate or reproducibility precision conditions [20].

The biological variation data critical appraisal checklist

When using BV data to define APS, the applicability of the APS depends on the BV data they are based on, being both reliable and relevant to the population the laboratory serves. In the consensus document from the EFLM Strategic Conference, it was stated that the user should “carefully assess the relevance and validity of the biological variation data” [10] used to set APS. On this background, the EFLM established in 2014 a Task Group for the biological variation

database, with the objective to appraise the quality of BV data that was publicly available [25]. One of the outcomes of this group's work, is the biological variation data critical appraisal checklist (BIVAC) [26], which is a tool for quality-assessing published BV studies. The BIVAC is designed to verify whether all essential elements that may influence upon veracity and utility of the associated BV estimates are included in a publication. The checklist consists of 14 quality items, which can be awarded scores A, B, C or D, indicating decreasing compliance. Critical quality items, such those referring to description of the study population, samples, sample material and timing of sampling, as well as the analytical method, may elicit a score D (unsuitable for use in clinical practice), if not appropriately performed or characterized in the study. The other quality items included in the checklist refer to necessary elements for interpretation and application of the BV data and may receive scores A, B and/or C. Since the publication of the BIVAC, an increasing number of studies adhere to this checklist, indicating they are fully BIVAC compliant. This illustrates how the BIVAC not only serves as a tool for appraisal of already published studies, but that it also may serve as a guide for those planning and publishing new studies.

The EFLM biological variation database and automatic calculation of BV-based APS

A second major initiative to make available high-quality BV data is the development of the EFLM biological variation database [2]. The EFLM biological variation database was launched in May 2019 and currently includes over 3060 BV data sets derived from 581 published and quality-assessed BV studies. Both details on the publications, the results for the 14 BIVAC quality items as well as a detailed BV minimum data set including around 30 descriptive items are registered for each publication included in the database. Additionally, BV estimates from studies with similar study design, performed in healthy adults and with BIVAC grade A, B and C, are included in meta-analyses to provide global BV estimates. Global BV estimates are presently published for 185 different measurands, with meta-analysis results for additional measurands in the line to be published. The automatically calculated APS can be found either by selecting the Meta-Analysis box on the front page or by searching for the measurand of interest (Figure 1). For the measurands for which there are sufficient quality BV data to perform a

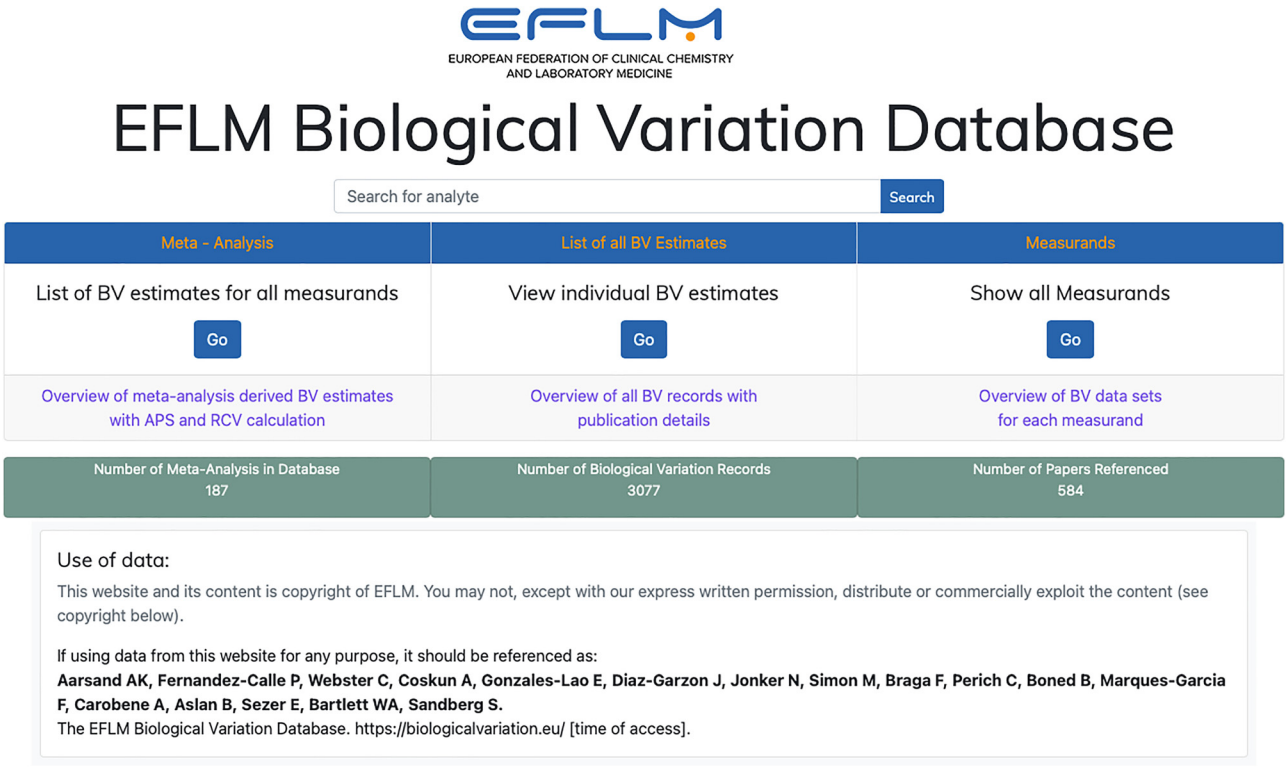

Figure 1: Screenshot of the front page of the online EFLM biological variation database. In this database, detailed information for biological variation data sets derived from critically appraised biological variation studies, as well global within-subject and between-subject biological variation estimates, derived from meta-analysis, are published and automatically updated whenever new data are added. For each measurand analytical performance specifications (APS) and reference change values (RCV) are also automatically calculated.

Table 1: Biological variation (BV) estimates for total calcium derived from meta-analysis, with associated analytical performance specifications (APS) for imprecision (CV_A), bias, allowable expanded measurement uncertainty (MAU) and total allowable error (TEa), as calculated in the EFLM biological variation database.

BV estimate (95 % CI)		Number of included BV datasets, n	Total number of BV data sets, n	APS	Imprecision (CV_A)	Bias	MAU	TEa
Within-subject BV (CV_I)	1.8 (1.7, 2.3)	20	65	Minimum	1.4	1.2	2.7	3.4
Between-subject BV (CV_G)	2.7 (1.1, 5.0)	17	58	Desirable	0.9	0.8	1.8	2.3
				Optimal	0.5	0.4	0.9	1.1

meta-analysis, users can also access automatically calculated APS, by selecting "Analytical Performance Specification". When accessing this feature, the database will automatically calculate and report the following BV based APS, i.e., imprecision, bias, TEa and MAU, defined for minimum, desirable and optimal, as illustrated, as an example, for total calcium in Table 1. In addition, data for measurands not fulfilling the criteria for the meta-analysis, or data from different states of health, age groups, gender and with other sampling intervals are also included in the database and may be of relevance to selected laboratories for setting APS.

Discussion

Setting appropriate APS is an important part of assuring the reporting of correct laboratory test results, and it is a continuously ongoing and at times challenging process. In this opinion paper, different approaches for setting APS based on BV have been presented. The APS described in this paper are also defined for different levels i.e., minimal, optimal and desirable (Table 1). However, the selection of the numerical factors defining these, may seem arbitrarily chosen [14]. In principle, only the APS for imprecision and MU are based solely on BV, whereas APS for bias and TEa might be considered a combination of BV and a type 1b approach.

When selecting APS, it must firstly be considered if the measurand in question is suited for the BV model. Furthermore, also the situations or scenario for which the APS are set must be considered. Are the APS to be used in the laboratory for method validation, lot to lot variation, internal quality control or for assessing participants in EQA schemes? Or is the APS an aspirational goal which the laboratory and manufacturers should strive to achieve in the future? One can argue that APS for imprecision and bias are more useful for a continuous evaluation of processes in the laboratory, whereas defining APS for MU may be particularly useful to be able to inform clinicians and patients about the uncertainty of laboratory results [21, 22].

Regardless of the selected APS approach, their applicability will depend on the BV data they are based on and whether these are relevant to the laboratory and type of population it serves. In today's BV database, global BV estimates are derived from studies in healthy adults where sampling has been from biweekly to monthly. For the most frequently requested measurands, many BV studies of acceptable quality may be available, whereas for others, only data from single studies are available. This underlines the importance for users assessing available BV data for its quality and relevance to the laboratory's populations prior to using such data to define their APS.

Conclusions and the way forward

There are different approaches for setting APS, but today, APS for many measurands are defined based on BV data. Which formulae to use is dependent on e.g. if the TE or the MU paradigm is used, the latter being metrologically the more correct. APS for bias is, however, appropriate for laboratories implementing TE paradigms. In the EFLM biological variation database, the APS that have been discussed in this paper are automatically calculated, accompanied by information on the strengths and limitations of the different approaches [2]. For the future, we will aim to add information on measurands for which APS based on BV should *not* be used and also to include APS for measurands where other models are preferable [12]. Furthermore, there is today a focus on data mining/"big data" studies to deliver BV data, which may facilitate delivery of BV data for new measurands and states of health. These studies are based on a different methodological approach than the classical prospective experimental BV studies. Work is ongoing to define a critical appraisal checklist applicable to such studies, allowing them

in the future to be included in meta-analysis in the database. There is also ongoing work to establish new concepts for setting appropriate APS based on BV, including how to better define minimum, desirable and optimum APS. Furthermore, paradigm shifts in metrology, including the draft VIM 4 will require even more extensive developments of APS. We will aim to include new APS in the database as they evolve, with the overall goal for the EFLM biological variation database to be an updated and valuable tool for laboratories, clinicians, and guidelines developers.

Acknowledgments: Professor Callum G. Fraser provided valuable input to this manuscript.
Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: The authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing interests: The authors state no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

1. 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.
2. Aarsand AK, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, Roraas T, et al. EFLM biological variation database [Internet]. https://biologicalvariation.eu [Accessed 29 Feb 2024].
3. Elevitch F. Proceedings of the 1976 aspen conference on analytical goals in clinical chemistry. In: Elevitch F, editor. Proceedings of the 1976 aspen conference on analytical goals in clinical chemistry. Illinois: College of American Pathologists; 1977.
4. Gowans EM, Petersen PH, Blaabjerg O, Hørder M. Analytical goals for the acceptance of common reference intervals for laboratories throughout a geographical area. Scand J Clin Lab Invest 1988;48:757–64.
5. Fraser CG, Petersen PH. Quality goals in external quality assessment are best based on biology. Scand J Clin Lab Invest 1993;53:8–9.
6. Petersen PH, Groth T, Verdier C-HD. 2. Principles for assessing analytical quality specifications ("AQSpecs") and their use in design of control systems. Ups J Med Sci 1993;98:195–214.
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;59:475–6.
8. Fraser CG, Petersen PH. Analytical performance characteristics should be judged against objective quality specifications. Clin Chem 1999;45:321–3.
9. Fraser CG, Petersen PH. Desirable standards for laboratory tests if they are to fulfill medical needs. Clin Chem 1993;39:1447–55.
10. 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.
11. Sandberg S, Røraas T, Aarsand AK. Biological variation and analytical performance specifications (Internet). In: Rifai N, Chiu RWK, Young I, Burnham C-AD, Wittver CT, editors. Tietz textbook of laboratory medicine, 7th ed. St Lous: Elsevier; 2022:335–56 pp.
12. 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.
13. Stockl D, Baadenhuijsen H, Fraser CG, Libeer JC, Petersen HP, Ricos C. Desirable routine analytical goals for quantities assayed in serum. Eur J Clin Chem Clin Biochem 1995;33:157–69.
14. Fraser CG, Petersen PH, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 1997;34:8–12.
15. International vocabulary of metrology–basic and general concepts and associated terms (VIM) (2008 version with minor corrections), Joint Committee for Guides in Metrology (JCGM), 3rd ed. Sevre, France: International Bureau of Weights and Measures; 2012.
16. Eisenhart C. Realistic evaluation of the precision and accuracy of instrument calibration systems. J Res Natl Bur Stand, Sect C 1963;67C:161.
17. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. Clin Chem 1974;20:825–33.
18. BIPM, IEC, IFCC, IUPAC, IUPAP, OIML. Guide to expression of uncertainty in measurement, 1st ed. Geneva: ISO; 1993.
19. Kallner A. Is the combination of trueness and precision in one expression meaningful? On the use of total error and uncertainty in clinical chemistry. Clin Chem Lab Med 2016;54:1291–7.
20. Coskun A, Theodorsson E, Oosterhuis WP, Sandberg S. Measurement uncertainty for practical use. Clin Chim Acta 2022;531:352–60.
21. Oosterhuis WP, Bayat H, Armbruster D, Coskun A, Freeman KP, Kallner A, et al. The use of error and uncertainty methods in the medical laboratory. Clin Chem Lab Med 2018;56:45–11.
22. Petersen PH, Stockl D, Westgard JO, Sandberg S, Linnet K, Thienpont L. Models for combining random and systematic errors. Assumptions and consequences for different models. Clin Chem Lab Med 2001;39:589–95.
23. Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. Clin Chem 2011;57:1334–6.
24. Fraser CG. Desirable performance standards for clinical chemistry tests. Adv Clin Chem 1983;23:299–339.
25. Panteghini M, Sandberg S. Defining analytical performance specifications 15 years after the Stockholm conference. Clin Chem Lab Med 2015;5:829–32.
26. Aarsand AK, Røraas T, Fernandez-Calle P, Ricos C, Díaz-Garzón J, Jonker N, et al. The biological variation data critical appraisal checklist: a standard for evaluating studies on biological variation. Clin Chem 2018;64:501–14.