EFLM Paper

Marith van Schrojenstein Lantman, Hikmet Can Çubukçu, Guilaine Boursier, Mauro Panteghini, Francisco A. Bernabeu-Andreu, Neda Milinkovic, Pika Mesko Brguljan, Solveig Linko, Duilio Brugnoli, Ruth O’Kelly, Christos Kroupis, Maria Lohmander, Luděk Šprongl, Florent Vanstapel and Marc Thelen*, on behalf of the European Federation of Clinical Chemistry, Laboratory Medicine EFLM Working Group Accreditation, ISO/CEN standards WG-A/ISO

An approach for determining allowable between reagent lot variation

<https://doi.org/10.1515/cclm-2022-0083>

Received February 1, 2022; accepted February 2, 2022;
published online February 16, 2022

Abstract: Clinicians trust medical laboratories to provide reliable results on which they rely for clinical decisions. Laboratories fulfil their responsibility for accurate and consistent results by utilizing an arsenal of approaches, ranging from validation and verification experiments to daily quality control procedures. All these procedures verify, on different moments, that the results of a certain examination procedure have analytical performance characteristics (APC) that meet analytical performance specifications (APS) set for a particular intended use. The APC can in part be determined by estimating the measurement uncertainty component under conditions of within-laboratory precision (u_{RW}), which comprises all components influencing the measurement uncertainty of random sources. To maintain the adequacy of their measurement procedures, laboratories need to distinguish aspects

that are manageable vs. those that are not. One of the aspects that may influence u_{RW} is the momentary significant bias caused by shifts in reagent and/or calibrator lots, which, when accepted or unnoticed, become a factor of the APC. In this paper, we postulate a model for allocating a part of allowable u_{RW} to between-reagent lot variation, based on the need for long-term consistency of the measurement variability for that specific measurand. The allocation manages the ratio between short-term and long-term variation and indicates laboratories when to reject or correct certain variations due to reagent lots.

Keywords: measurement uncertainty; measurement procedures; position paper; reagent-lot.

Introduction

The results of laboratory measurements are a key element for clinical decision-making in medicine. Reliability and

*Corresponding author: Marc Thelen, Result Laboratory for Clinical Chemistry, Amphia Hospital, Breda, The Netherlands; and Department of Laboratory Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands, E-mail: MThelen@amphia.nl

Marith van Schrojenstein Lantman, Result Laboratory for Clinical Chemistry, Amphia Hospital, Breda, The Netherlands; and Department of Laboratory Medicine, Radboud University Medical Centre, Nijmegen, The Netherlands. <https://orcid.org/0000-0002-5454-990X>

Hikmet Can Çubukçu, Interdisciplinary Stem Cells and Regenerative Medicine, Ankara University Stem Cell Institute, Ankara, Turkey. <https://orcid.org/0000-0001-5321-9354>

Guilaine Boursier, Dept of Genetics, Rare Diseases and Personalized Medicine Rare Diseases and Autoinflammatory Unit, CHU Montpellier, Univ Montpellier, Montpellier, France

Mauro Panteghini, Department of Biomedical and Clinical Sciences “Luigi Sacco”, and Research Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milano, Italy

Francisco A. Bernabeu-Andreu, Servicio Bioquímica - Análisis Clínicos, Hospital Universitario Puerta de Hierro, Madrid, Spain

Neda Milinkovic, Department of Medical Biochemistry, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia

Pika Mesko Brguljan, Department of Clinical Chemistry, University Clinic for Respiratory and Allergic Diseases, Golnik, Slovenia

Solveig Linko, Linko Q-Solutions Helsinki, Finland

Duilio Brugnoli, Clinical Chemistry Laboratory, Spedali Civili, Brescia, Italy. <https://orcid.org/0000-0001-9420-9961>

Ruth O’Kelly, Association of Clinical Biochemists in Ireland, Dublin, Ireland

Christos Kroupis, Department of Clinical Biochemistry, Medical School, Attikon University General Hospital, National and Kapodistrian University of Athens, Haidari, Greece

Maria Lohmander, Regional Laboratoriemedicin, Sahlgrenska Universitetssjukhuset, Trollhättan, Sweden

Luděk Šprongl, Clinical Laboratory, Hospital Kladno, Kladno, Czech Republic

Florent Vanstapel, Belgium and Department of Public Health, Laboratory Medicine, University Hospital Leuven, Biomedical Sciences Group, Leuven, Belgium

therefore constraint of measurement uncertainty (MU) of laboratory results is therefore of vital importance as to not affect clinical decision-making in a negative way. The MU of patient results is made up of various components of analytical variation, some of which have a time-based frequency, like hourly variation and day-to-day variation, while others follow an event-based pattern, like calibration variation, reagent lot variation or calibrator variation caused by value assignment [1].

Monitoring MU with internal quality control (QC) is central to maintain the analytical performance characteristics (APC) within the analytical performance specifications (APS), which is essential for patient results that are fit for the intended use [2]. According to document ISO/TS 20914:2019, MU is made up from the uncertainty of the calibrator and of higher-order reference, if applicable, the uncertainty of the bias correction, and the within-laboratory uncertainty (u_{RW}). ISO/TS 20914:2019 defines u_{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 both short-term variation and intermediate precision as routine changes to measuring conditions) [1]. Examples of these measuring conditions are lot changes of reagents and calibrators, and instrument maintenance, and each of these variables may contribute to the MU of clinical samples [1, 3].

Unfortunately, no indications exist for the division of acceptable variation between its typical sources of within-lot and between-lot, and between-instrument variation. The element of between-instrument variation can be managed by the virtual analyzer concept where all analyzers have to conform to the same mean and SD set for internal QC, but for the division between within- and between-reagent lot variations (u_{brlot}) a new theoretical, but also practical concept is needed [4].

Current practice and shortcomings

In current practice, subsequent reagent lots are usually compared against each other [5]. The analysis of subsequent lots assumes that with a new lot there is a possible persistent realization of lot-associated random error. The Clinical and Laboratory Standards Institute (CLSI) EP26A gives an extensive description of this approach, using patient samples to compare subsequent lots [6]. Additionally, EP-31-A-IR provides guidance on how laboratories can verify comparability of individual patient results within a health care system [7].

While this approach enables the consistency of results of individual patients across two subsequent lots, there are

some major concerns that should be considered. Firstly, the assessment of a current lot change can only validate the sufficiency of closeness for two lots. Therefore, in a single-site setting, every lot is only compared twice in its lifetime, i.e., with its predecessor and its successor. These analyses aim to guarantee a smooth transition towards the next lot and minimize direct risk for patients due to change of two lots. However, this approach lacks criteria for all lot variation, thereby overlooking, and allowing the introduction of long-term trends or other cumulative effects that may exceed the long-term lot-to-lot criteria. EP26A relies on other mechanisms, such as external quality assessment (EQA), to detect such trends. If unacceptable trends are detected not earlier than by EQA interpretation, results produced over an extended period might need to be recalled, probably long after medical decision may have been based on unacceptable results.

Secondly, the EP26A approach requires commutable samples for the assessment of reagent lot variation, which may pose difficulty for the laboratory. While it is postulated that patient samples or, alternatively, commutable QC samples are required to prevent any non-commutability issues in commercial control materials, they can be difficult to obtain at the right time, and the right sample size [8]. In addition, appropriate storage and stability over the storage period is necessary, which may add to the complexity when carrying out the EP26A protocol.

Debate has been going on whether using internal and/or external QC materials may be a valid alternative when assessing reagent lot variation, at least for common biochemistry tests [9–11]. A recent study of Kim et al. [12] reported that in their setting the commutability of a commercial QC material was good enough for the lot-lot variation study of 61 general biochemistry and urine tests, but not satisfying for at least 4 of 50 immunoassays in their study [12]. For the purpose of the study of between-lot variation, the materials used need to be commutable between reagent-lots, within a used method. Although commercial control materials are known for their lack of between-method commutability, in most methods they may prove to have satisfactory within-method between-lot commutability. For immunoassays however, the chances are high that laboratories need to select and apply other materials when commercial controls prove to lack within method between-lot commutability.

Thirdly, the EP26A reports that for some measurands, a large sample size can be needed to obtain enough statistical power [8]. However, a large sample size can be hard to obtain, and, for measurands where the material is difficult to obtain, this may result in an unfeasible protocol [8, 12]. Pooling samples may be an option to obtain a large sample

size, but the process of pooling may cause artifacts. Another disadvantage of pooling samples is that there is a lack of concentration diversity unless samples are pooled in well-chosen concentration ranges.

Lastly, the approach of comparing subsequent lots always originates from the principle that the former lot is the one that should be regressed towards. At the level of the first lot, there is no lot to compare to, and therefore this lot is often deemed as a “true” or “typical” one. However, if acceptability of the first lot-bias is not validated, chances are that this lot could perform atypically. Different laboratories could have different first and true lots of the same product, thereby making every comparison problematic.

In summary, when using the EP26A protocol for the study of between-lot variation, the above reported concerns are often valid, and this necessitates efforts to develop a new approach to assess the impact of the reagent lot variation.

Clinical relevance

The relevance of between-lot variation may be different for different measurands. The impact of between-lot variation is smaller for measurands that are primarily used for diagnosis in a short-term period rather than long-term monitoring. For example, cardiac troponin in the diagnosis of non-ST-elevation myocardial infarction in individual patients barely suffers from between-lot variation, because the recommended serial measurements are executed using the same reagent lot in a brief time span. In contrast, the long-term monitoring of measurands such as prostate-specific antigen (PSA) or glycated hemoglobin (HbA_{1c}) is fundamental in the management of prostate cancer and diabetes mellitus, respectively. In these cases, a significant reagent lot variation may be mistaken for a clinical change by the clinicians and patients may be treated as such.

The impact of measurement frequency and timing is shown in Figures 1, where the difference in amount of

patient samples per lot influences the ability and the rapidity to detect an aberrant reagent lot. In Figure 1A the patient appears to have a larger variation between consecutive samples, which is partially prompted by the reagent lot change. As the number of samples for an individual patient within a reagent lot rise (Figure 1B), the observed variance by the clinician in patient results is uncommonly caused by a reagent lot change, and therefore, a reagent lot change can mirror a clinical change. Examples of this were reported in literature for a change of 5.5 mmol/mol in HbA_{1c} [13], a change in cerebrospinal fluid biomarkers Aβ1-42 and p-tau, which led to alternative Alzheimers’ disease classification [14], false positive results with anti-HIV antibody assays [15], and a false increase in insulin-like growth factor 1 results [16].

Moving patient averages can be used to monitor long-term assay stability [13, 16]. However, there are some concerns against this approach, because of seasonal differences (which may be solved by taking appropriate time-frames into consideration and being acquainted with the expectable seasonal variance) and filters used in the statistical treatment of the data [8]. Liu et al. proposed an interesting approach where EP9 results are used [17], but laboratories may argue that such approach is unfeasible and unneeded to deploy for every reagent lot change. These shortcomings are endorsed by several publications, and other ways of tackling reagent lot variation have been proposed; however, these approaches have been unsuccessful in shifting laboratories worldwide to their use [8, 13, 16, 17]. The main obstacle for adopting these methodologies is the required statistical computation, which may be complicated if laboratories do not have adequate software(s).

The simulation reported in Figure 1 shows that there is a need to assess reagent lot variation to avoid clinical implications. Therefore, a guidance is needed on how to check reagent lot-to-lot variation in a statistically sound and practical way.

Figure 1: Differences in impact of lot variation over time. Blue dots represent internal QC samples, red dots represent patient samples, and vertical dotted line represents a change in reagent lot. In (A) a patient is measured once per reagent lot, where in graph (B) the patient is measured six times per lot.

When does a reagent lot change become apparent?

The effect of reagent lot variation becomes apparent when a lot has changed, and the laboratory reports the first measurements using the new lot. The results of the new lot become apparent as “different” when they are recognized as different from the previous reported results by the clinician who requested and interpret them. This is dependent on two factors: (1) the ratio of within-lot and between-lot variation, and (2) the typical measurement frequency of an individual within an average length that a lot is in use, which is often determined by policies of suppliers and/or laboratories.

1) The ratio between within-lot and between-lot variation

A reagent lot change becomes apparent when the persistent random error due to the between-lot variation is larger than the usual within-lot variation. Nowadays, this has become more apparent due to achievements in minimizing within-lot variation by *in vitro* diagnostics (IVD) manufacturers, resulting in similar within-lot variation for different lots. Whereas within-lot variation is mostly dependent on mechanical and optical instrument imprecision, between-lot variation is caused by the manufacturing process of reagents and calibrators, and, in the case of calibrators, by their subsequent value assignment. Although manufacturers are formally required to adhere to an auditable traceability chain, and recent development of ISO 17511:2020 on metrological traceability in laboratory medicine makes these requirements even sharper [18], these chains are often longer than necessary and may use calibrator materials with uncertain commutability resulting unnecessary impact on the MU. Therefore, the ratio of within-lot and between-lot variation can indicate, although indirectly, the likelihood of incorrect medical decision-making based on lot-effects.

2) The typical measurement frequency of an individual in a typical period of lot use

The relevance of reagent lot differences is dependent on the ability of the requesting clinician to form a customary feeling of the patient variation. A feeling of “usual variance” aids clinicians in recognizing patterns in patients which warrant clinical decisions. In addition, patients track their own laboratory results, where a sudden change in outcome may cause stress and anxiety. In practice, the “usual variance” in

a stable patient is a combination of within-patient biological variation and analytical variation, which can either be within-reagent lot or a combination of within- and between-reagent lot variation (as shown in Figure 1).

In this case, the number of samples that a clinician has requested within the same reagent-lot period coincides with his/her ability to detect a between-lot change (and mistake it for clinical patient change). If the number of observations for one patient using one reagent lot increases, the anticipated measurement variation of the requesting clinician will be based on that typically experienced with within-lot variation. This results in an increased ability to distinguish the difference between lots and, therefore, the ability to experience between-lot changes becomes more likely when more samples of the same patient are analyzed using the same lot.

Mathematical evaluation of allowable between-reagent lot variation

There are various sources of MU that occur during the analytical phase. A graphical simulation is outlined in Figure 2. These sources of MU add up to comprise the u_{RW} , which can be estimated in accordance to ISO/TS 20914:2019, Section 5.3 [1, 19]. Statistically, the uncertainty component under conditions of within-laboratory precision (u_{RW}) is made up by the square root of the quadratic sums of individual uncertainties. If a division is done of u_{RW} towards within-reagent lot variation (u_{wrlot}) and between-reagent lot variation (u_{brlot}), the formula of Eq. (1) is obtained.

Figure 2: Graphical representation of sources of variation, and how they add up to comprise the total variation.

Figure 3: Example of distribution of internal quality control samples as a function of reagent lot.

A and C=maximum allowable between-reagent lot variation, B=average of all lots, blue dots=internal QC results, E=internal QC result mean of lot 1, F=+2SD limit of lot 1, G=-2SD limit of lot 1, H=distribution of lot means.

$$u_{Rw} = \sqrt{u_{wrlot}^2 + u_{brlot}^2} \quad (1)$$

During the use of a reagent lot, internal QC materials are measured (shown in Figure 3 as blue dots). Within a lot, internal QC results have a mean (blue horizontal line) and an SD (u_{wrlot}). In Figure 3 the variation is shown as a 95% confidence interval, depicted as vertical lines with small horizontal lines at the ends.

The between-lot variation (u_{brlot}) is obtained by assessing the variance between the internal QC means of the different reagent lots. In Figure 3 this is depicted as curve H. The blue band in Figure 3 depicts the ± 2 SD range of variability in which different reagent lots reside.

A reagent lot change would become apparent when the 95% confidence intervals of measurements within two lots would not overlap. This is represented in Figure 3 where the 95% confidence intervals of lot 1 and lot 8 do not overlap. Lot 8 may be acceptable if it would have been the second one used, with only lot 7 as a reference, but with the knowledge of lots 1–7, lot 8 becomes atypical. Therefore, we equalize the maximum allowable between-lot variation (Eq. (2a)) with the 95% confidence interval of the within-lot variation (Eq. (2b)).

$$\text{between lot range} \approx \pm 1.96 \times u_{brlot} \quad (2a)$$

$$\text{between lot range} \leq \pm 1.96 \times \frac{u_{wrlot}}{\sqrt{n}} \quad (2b)$$

In the calculation of this confidence interval, n is taken for the number of measurements of a patient within a single reagent lot. We use n as the number of patient measurements in accordance with the analytical variation as obtained by internal QC to obtain a measure for the standard error of the mean (SEM). Usually, the SEM is used to describe the uncertainty of the mean value, but in our case, we propose to use this to describe the uncertainty of the mean patient value within the timeframe of one reagent lot

use. We think this calculation is an appropriate reflection of the degree of habituation a doctor has about a patient status. The maximum value of n is arbitrarily set at 6 to enable a minimum budget to be available for between-reagent lot variation. This results in Eq. (3) for the allowable between-lot uncertainty:

$$u_{brlot} \leq \frac{u_{wrlot}}{\sqrt{n}} \quad (3)$$

A graphical representation of the allowable within- and between-lot variation is depicted in Figure 4.

The allowable variation for u_{brlot} is derived by using its relationship with u_{wrlot} as depicted in Eq. (3) to obtain a relative relationship with the APS set for the measurand in question.

We follow ISO/TS 20914:2019 to define the APS for the allowable MU of patient results $u(y)$ in terms of the u_{Rw} and MU of the calibrator (u_{cal}) [1]. The relationship between APS, u_{cal} and the u_{wrlot} and u_{brlot} components of u_{Rw} is shown below. As we evaluate the allowable between-reagent lot variation by dividing MU, we base our APS on $u(y)$, as we note that for this calculation sources of MU on the higher steps of the calibration hierarchy should be included. We have therefore included u_{cal} in the equation below so that, employed APS for u_{Rw} , as reported in the example above, may be a fraction of the APS applied for $u(y)$ [20]. By choosing APS based on $u(y)$, the formula Eq. (4) applies and allows for the expression of allowable $u_{brlot,a}$ in terms of APS.

Figure 4: Division of allowable measurement uncertainty contributions of within-lot variation (u_{wrlot}) and between-lot variation (u_{brlot}), based on the amount of single patient measurements within a reagent lot.

$$\begin{aligned}
 APS &\geq \sqrt{u_{Rw}^2 + u_{cal}^2} = \sqrt{u_{wrlot,a}^2 + u_{brlot,a}^2 + u_{cal}^2} \\
 APS^2 &\geq (u_{brlot,a} \times \sqrt{n})^2 + u_{brlot,a}^2 + u_{cal}^2 \\
 APS^2 - u_{cal}^2 &\geq u_{brlot,a} \times \sqrt{n+1} \\
 u_{brlot,a} &\leq \frac{\sqrt{APS^2 - u_{cal}^2}}{\sqrt{n+1}} \quad (4)
 \end{aligned}$$

If the MU of the calibrator is unknown, Eq. (4) can be simplified to the formula below.

$$u_{brlot,a} \leq \frac{APS}{\sqrt{n+1}}$$

Using biological variation model as the source of the APS for $u(y)$

The biological variation model states that the allowable MU budget, set at desirable quality level, should be less than half of the within-subject biological variation [21]. We have applied the “rule” of $\leq 0.5 \times CV_I$ as a way to define the APS set for $u(y)$ in Eq. (4), which results in Eq. (5).

$$u_{brlot,a} \leq \frac{\sqrt{(0.5 \times CV_I)^2 - u_{cal}^2}}{\sqrt{n+1}} \quad (5)$$

With Eq. (5) we can determine the allowable between-lot variation as a product of the within-subject biological variation and the number of observations within a lot. For measurands that are unable to perform in accordance with the required APS by biological variation, guidelines of (inter)national scientific societies can set a short-term APS at an achievable level.

How to derive allowable between-lot variation

- (1) Establish the within-subject biological variation (CV_I) (an updated source of information is provided by <https://biologicalvariation.eu/>, [22]),
- (2) Establish n , as the typical number of results of a single patient within the reagent lot use period that is based on the intended use or professional opinion. These data may be derived from the laboratory information system. This number will be typically higher for measurands used in daily clinical practice (i.e., plasma creatinine and glucose) than for measurands used for the long-term management in an outpatient setting (i.e., HbA_{1c} or PSA),
- (3) Use Eq. (4) to determine if between-reagent lot variation complies with the allowable MU,

Use the mean of the reagent lots $\pm\% u_{brlot}$ to establish boundaries for new reagent lots.

A few worked-out examples are listed in Supplemental Material 1.

Benefits

Implementation of the suggested approach enables systematic monitoring the amount of variation derived from reagent lot changes. This approach would reduce the strain on laboratories for verification testing when a new lot is introduced, as only the compliance of the new lot with the distribution of all lots must be tested. Furthermore, this approach can manage long-term drifts in reagent lot means and hence controls the between-reagent lot variation.

Discussion

There is a potential significant effect of reagent lot variation on patient care. Reagent lot variation may introduce a temporary bias into the measuring system and, if unnoticed, included in the long-term imprecision estimate. If the reagent lot variation is mistaken for clinical change, consequences can be serious. Therefore, it is of vital importance to assess reagent lot variation in the laboratory practice, testing its effects, and determining the risk for patients. We do, however, note that whilst the between-reagent lot variation may affect the quality of laboratory analytical results, the quality of laboratory *information* is a much broader and complex issue.

Our proposed approach is applicable for measurands for which APS can be defined for a given intended use and MU can be calculated from available data, but does not solve the problem a laboratory is facing with the introduction of a new measurement procedure. In this case, there is a limited knowledge about the within- and between-lot variations that will appear evident after a prolonged use by the laboratory. To overcome this problem, manufacturers should provide typical performance information or, as an alternative, the user community of the same IVD device can share such information. Some manufacturers of third-party QC materials provide between-laboratory information. Within the European Union, guidance on method validation and on post-market surveillance could nudge the manufacturers to compile this information over the lifecycle of their devices. EQA organizers can also take a role in collecting and distributing information regarding reagent lot variation, as this information can be noticed in EQA results [9, 23]. However, to achieve this, EQA providers should use a category 1 scheme using commutable samples [24]. Working out different options for the provision of between-lot variation information will identify a lot of obstacles to be resolved and is beyond the scope of the present paper.

There are certain assumptions using our approach. First, it is based on the conviction that variation in time and variation in place should be treated equally. This statement infers for example that allowable variation within one instrument should be treated equal to the allowable variation between instruments. We induce this statement from the perspective of the patient: the total variation subjected to patient results should be maintained, irrespective of its source. In other words, the difference between obtained results of the patient between measurements done today vs. tomorrow should be similar to the difference between the obtained result of analyser X vs. analyser Y on the same day. This perspective does not give more weight to certain sources of variation. Laboratories implement this principle via the institution of the virtual analyser [25]. With the choice for similar acceptance criteria for variation in time as in place, we do not want to suggest that laboratories should ignore the difference in causes of variation. Only by knowing the source of variation, the variation can be managed.

Another assumption of this model is that IVD manufacturers compare the performance of each new reagent batch to a higher-order reference material and/or method in the selected metrological traceability chain [26]. Although the ISO 17511:2020 standard requires such an approach, harmonization of this standard with IVD and FDA regulations is still ongoing [18, 27]. Therefore, it is currently unclear to which degree IVD manufacturers adopt the mentioned approach. Lastly, it may be problematic to find an appropriate representative count of measurements, as a study assessing the frequency of laboratory testing on medical decisions claimed that it would not be possible to determine an average frequency with which a patient is tested [28]. However, authors did not report any evident impossibility to assess the frequency of a single laboratory test [28]. In addition, this study did not take the impact of intended use of the different measurands into account, which is an essential part of our approach. We do, however, recognize the difficulty in finding a representative count, as the average frequency of laboratory tests may be dependent on age, sex, diagnosis, hospital, medical specialty, sociodemographic correlates and physician characteristics [29–32]. Therefore, we endorse proper determination of the average frequency of testing.

A critical note of this paper is the used material to assess the allowable between-reagent lot variation. Firstly, we have postulated that, although ISO/TS 20914:2019 [1] supports the use of commercial control materials to study cross-lot uncertainty, its fitness for between-reagent lot studies needs verification of its between-reagent lot commutability. For immunochemistry tests and other assays where internal QC targets differ between reagent lots, our suggested approach first needs the selection and application of commutable

materials. Secondly, there are restraints in terms of the control materials that can be used to assess the between-reagent lot variation, the same criteria that apply to materials used to derive the MU of the analytical system. We advise using matrixed material from independent sources that closely resemble authentic patient samples with relevant concentrations for the intended use of the measurand. Our aim was to describe an example of a realistic approach that can be accomplished with currently available procedures and materials that are already in use [33].

When users rely on laboratory measurements, they want to see similarity of analytical performance across lots of reagents and calibrators, instruments, time, and locations. With this paper, we hope to focus laboratory communities' attention on the reagent lot variation, and how it should relate to the total variation, and on its effects on patient care. We believe that IVD manufacturers have the responsibility to communicate to end-users the amount of between-lot variation that laboratories can anticipate. The EU IVD Regulation 2017/746 has translated this responsibility into a legal obligation, specifically, ANNEX XIII describes "performance evaluation" as a documented continuous process confirming the performance of the devices put on the market throughout their lifetime [27]. Laboratory professionals have the responsibility to check whether the (anticipated) between-lot variation meets the medical need in the intended use of laboratory measurements. Such requirements can differ between different settings and depend on both allowable total MU budget and typical number of measurements for a patient within the typical lot usage period. Laboratory professionals can apply such knowledge in the selection process of new measuring systems. For those already in use, the information can be used to demonstrate if a bias due to the lot variation becomes larger than allowable.

Conclusions

With the use of the $u_{\text{brlot},a} \leq \frac{\sqrt{\text{APS}^2 - u_{\text{cal}}^2}}{\sqrt{n+1}}$ laboratories can assess the partition of the MU budget among the within- and the between-reagent lot variation, assuring that patient results do not suffer from clinically significant between-lot variation. This approach is a first step to answer the call by Loh et al. for a risk-based solution that does right to the clinical impact of between-reagent lot variation [34].

Research funding: None declared.

Author contribution: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

Informed consent: Not applicable.

Ethical approval: Not applicable.

References

- International Organization for Standardization. Medical laboratories – Practical guidance for the estimation of measurement uncertainty. ISO/TS 20914:2019; 2019.
- Infusino I, Panteghini M. Measurement uncertainty: friend or foe? Clin Biochem 2018;57:3–6.
- Braga F, Pasqualetti S, Aloisio E, Panteghini M. The internal quality control in the traceability era. Clin Chem Lab Med 2020;59:291–300.
- Schoenmakers CHH, Naus AJM, Vermeer HJ, van Loon D, Steen G. Practical application of sigma metrics QC procedures in clinical chemistry. Clin Chem Lab Med 2011;49:1837–43.
- Katzman BM, Ness KM, Algeciras-Schimmich A. Evaluation of the CLSI EP26-A protocol for detection of reagent lot-to-lot differences. Clin Biochem 2017;50:768–71.
- CLSI. User evaluation of between-reagent lot variation. CLSI guideline EP26-A. Wayne, PA: Clinical and Laboratory Standards Institute; 2013.
- CLSI. Verification of Comparability of Patient Results Within One Health Care System. CLSI guideline EP31-A-IR. Wayne, PA: Clinical and Laboratory Standards Institute; 2012.
- Thompson S, Cheshier D. Lot-to-lot variation. Clin Biochem Rev 2018;39:51–60.
- Kristensen GBB, Meijer P. Interpretation of EQA results and EQA-based trouble shooting. Biochem Med 2017;27:49–62.
- Don-Wauchope AC. Lot change for reagents and calibrators. Clin Biochem 2016;49:1211–2.
- Miller W, Ere A, Cunningham TD, Oladipo O, Scott MG, Johnson RE. Commutability limitations influence quality control results with different reagent lots. Clin Chem 2011;57:76–83.
- Kim S, Chang J, Kim S-K, Park S, Huh J, Jeong T-D. Sample size and rejection limits for detecting reagent lot variability: analysis of the applicability of the Clinical and Laboratory Standards Institute (CLSI) EP26-A protocol to real-world clinical chemistry data. Clin Chem Lab Med 2020;59:127–38.
- Thaler MA, Iakoubov R, Bietenbeck A, Luppa PB. Clinically relevant lot-to-lot reagent difference in a commercial immunoturbidimetric assay for glycated hemoglobin A1c. Clin Biochem 2015;48:1167–70.
- Vos SJB, Visser PJ, Verhey F, Aalten P, Knol D, Ramakers I, et al. Variability of CSF Alzheimer's disease biomarkers: implications for clinical practice. PLoS One 2014;9: e100784.
- Sasaki M, Sasaki K. Significance of the control serum measurement in infectious disease tests – a case of lot-to-lot variation of anti-HIV antibody assay kit. Rinsho Byori 2002;50:717–20.
- Algeciras-Schimmich A, Bruns D, Boyd J, Bryant S, Fortune K, Grebe S. Failure of current laboratory protocols to detect lot-to-lot reagent differences: findings and possible solutions. Clin Chem 2013;59:1187–94.
- Liu J, Tan CH, Loh TP, Badrick T. Detecting long-term drift in reagent lots. Clin Chem 2015;61:1292–8.
- International Organization for Standardization. In vitro diagnostic medical devices – Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples. ISO 17511:2020.
- Braga F, Panteghini M. The utility of measurement uncertainty in medical laboratories. Clin Chem Lab Med 2020;58:1407–13.
- Braga F, Panteghini M. Performance specifications for measurement uncertainty of common biochemical according to Milan models. Clin Chem Lab Med 2021;59:1362–8.
- Cotlove E, Harris EK, Williams GZ. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. 3. Physiological and medical implications. Clin Chem 1970;16:1028–32.
- EFLM Biological Variation. Available from: <https://biologicalvariation.eu/> [Accessed 1 Dec 2021].
- Kristensen GBB, Christensen NG, Thue G, Sandberg S. Between-lot variation in external quality assessment of glucose: clinical importance and effect on participant performance evaluation. Clin Chem 2005;51:1632–6.
- Jansen RTP, Cobbaert CM, Weykamp C, Thelen M. The quest for equivalence of test results: the pilgrimage of the Dutch Calibration 2.000 program for metrological traceability. Clin Chem Lab Med 2018;56:1673–84.
- Calleja J. Parallel processing and maintaining adequate alignment between instruments and methods. Clin Biochem Rev 2008;29:S71–7.
- Miller WG, Greenberg N. Harmonization and standardization: where are we now? J Appl Lab Med 2021;6:510–21.
- Regulation (EU) 2017/746 of the European parliament and of the council of 5 april 2017 on in vitro diagnostic medical devices and repealing directive 98/79/EC and commission decision 2010/227/EU (text with EEA relevance). 2017. Available from: <http://data.europa.eu/eli/reg/2017/746/oj/eng> [Accessed 2 Aug 2021].
- Ngo A, Gandhi P, Miller WG. Frequency that laboratory tests influence medical decisions. J Appl Lab Med 2017;1:410–4.
- Georgiou A, Sezgin G, Li L, Wilson R, McCaughey EJ, Lindeman R, et al. Who gets a laboratory test in hospital, why, and how often? A retrospective observational study of 4 Australian hospitals. J Appl Lab Med 2019;4:343–54.
- Barber J, Guo M, Nguyen LT, Thomas R, Turin TC, Vaska M, et al. Sociodemographic correlates of clinical laboratory test expenditures in a major Canadian city. Am J Clin Pathol 2017; 148:91–6.
- Hardwick DF. Pathology—medicine's search for meaning in the new millennium. Hum Pathol 2002;33:1.
- Vacellio E. Australian Government Department of Health. Examination of variation in hospital pathology investigations by diagnosis-related groups and associations with outcomes and costs: report to Commonwealth of Australia. Sydney: Department of Health, Quality Use Pathology Committee. Report to Commonwealth of Australia, Department of Health, Quality Use of Pathology Committee. Australian Institute of Health Innovation, Macquarie University; 2015.
- Braga F, Infusino I, Panteghini M. Performance criteria for combined uncertainty budget in the implementation of metrological traceability. Clin Chem Lab Med CCLM 2015;53: 905–12.
- Loh TP, Sandberg S, Horvath A. Lot-to-lot reagent verification: challenges and possible solutions. Clin Chem Lab Med 2022;60: 675–80.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/cclm-2022-0083>).