EQA- derived metrics to assess overall instrument performance (Letter)

Author

Badrick, Tony, St John, Andrew

Published

2016

Journal Title

Clinical Chemistry and Laboratory Medicine (CCLM)

Version

Version of Record (VoR)

DOI

10.1515/cclm-2015-0680

Letter to the Editor

Tony Badrick* and Andrew St John

EQA-derived metrics to assess overall instrument performance

DOI 10.1515/cclm-2015-0680
Received July 16, 2015; accepted November 11, 2015; previously published online December 17, 2015

Keywords: laboratory management; method comparison; quality assurance.

To the Editor,

Recently sigma metrics have been used to assess assay fitness for purpose using external quality assessment surveys [1, 2]. One of the major decisions facing clinical laboratories is the evaluation and purchase of new equipment. This exercise is a considerable one with the costs of both the initial equipment and ongoing consumable costs having a significant impact on the laboratory budget. Many factors may be taken into account when comparing analysers including hardware and consumable costs, environmental costs (power and water), ongoing training and maintenance support. But perhaps most attention is devoted to the analytical performance of individual assays since poor performance can have deleterious effects on reagent and human resource costs (repeats, recalibration), staff morale, and customer satisfaction in terms of poor turnaround times and inappropriate clinical decisions which can reduce laboratory credibility.

While the performance of individual assays can be assessed through standard evaluation processes, it is difficult to compare instruments overall. When assessing many different measurands, the reality is that no one manufacturer or instrument type has all the best performing assays. Thus with a broad range of assays on board the instrument (>30), while all the commonly requested measurands may perform well, some low volume assays could deliver significant performance issues that may increase the sorts of costs mentioned above. This leads to purchase decisions being based on a few major measurands or by referring to other users of the instrument being evaluated. Consequently there is a need to have a more evidence-based assessment process that takes all measurands into consideration.

All laboratories are involved in external quality assurance programs (EQA) and summaries of instrument and method performance for all measurands are available as soon as new instruments or methods are being routinely used. This data provides independently collected data that can serve as the input to a metric which can provide an overall assessment.

A metric for each measurand can be calculated as $m_i=CV_{inst}/TEa_{inst}$, where CV_{inst} is the CV of the instrument group obtained from the EQA for that particular measurand (i) and analyser, and TEa_{inst} is the total allowable error (TEa) for the measurand. The TEa for that particular measurand determined by the laboratory and may be based on biological variation, regulatory requirements or clinical judgment; in Australia this is termed the allowable limit of performance by the Royal College of Pathologists of Australasia Quality Assurance Program (RCPAQAP) [3]. In the case of the RCPAQAP data the CVs represent the performance over a three month cycle of all users of a particular analyser. The CVs are also based on the fit of a linear regression to all results obtained from all the EQA samples. Thus the CVs are based on the standard error of the estimate over the median of values [4].

This metric has been previously used by Fraser [5] to set quality specifications for assays. He suggested that desirable performance is met when $m_i<0.5$; optimal performance is met when $m_i<0.25$; and minimum performance is met if $m_i<0.75$. The metric is the inverse of the capability, which is TEa/CV, without bias included. The CV is the CV of the method group determined from the EQA data, which represents the performance of the assay in the field. We have not taken bias into account in the calculation as this is difficult to objectively estimate and imprecision is the major concern with QC processes. Furthermore, bias can only be estimated if the QA/QC material used by the laboratory is commutable and traceable to a reference material. As there are multiple measurands on

***Corresponding author: Tony Badrick,** RCPAQAP, St Leonards, Sydney, NSW, Australia, E-mail: Tony.badrick@rcpaqap.com.au
Andrew St John: ARC Consulting, Perth, WA, Australia

Table 1: Data from EQA Survey for specific measurands with calculated $m=CV_{int}/TEa_{int}$ for two analysers.

Measurand	Metric ($m_i=CV_{inst}/TEa_{inst}$)	
	Analyser A	Analyser B
Urea	0.5	0.66
Sodium	0.46	0.36
Potassium	0.5	0.35
Glucose	0.75	0.9
Creatinine	0.62	0.18
Bilirubin	1.3	0.8
Bicarbonate	0.45	0.48
Alkaline phosphatase	1.18	0.68
Albumin	0.74	0.44
ALT	0.64	0.68
$M=\sum m_i/n$	0.714	0.711

a platform to consider we take the average of these to form an average performance metric $M=\sum m_i/n$, where n is the total number of measurands on the platform.

As an example of the use of this metric consider the following data obtained from the RCPAQAP End of Cycle Report [6] for two major clinical analysers. For each of a range of common measurands, based on the performance of users throughout the Cycle (4 months), we calculated a performance metric m_i for each measurand on each analyser based on the CV of that test (Table 1).

As can be seen some assays perform better on either one or the other analyser. If we calculate the metric $M=\sum m_i/n$ we obtain values of 0.714 for Analyser A and 0.711 for Analyser B so, over all there is little difference.

Next we consider using a weighting factor to derive a modified metric $M=\sum f_i^* m_i$; the weighting f_i^* could depend on the importance of the measurand on that analyser in that laboratory so it could be set at 1 for the high volume measurands (sodium, potassium, creatinine), 0.5 for lower volume measurands (albumin, cholesterol) and 0.25 for very low volume assays (lipase). Alternatively the factors could be based on urgency of the measurand (4 for potassium, 2 for AST and 1 for triglyceride) or the factor could be set on their potential clinical importance. The weightings would be purpose deigned and varied by the laboratory but their value is not critical as long as they are used consistently for all instruments being evaluated. Thus the total metric would therefore become $\sum f_i^* m_i$. We still need to take an average so we divide the sum of performance by the total weighting $\sum n_i w_i$ where w_i is the weighting given for the measurand n_i.

If we now add a weight to those measurands which may be of more significance we see a different picture. Let us apply a weighting of 2 to potassium, glucose and creatinine rather than the same weight of 1 which had been applied to all measurands equally previously. If we now calculate the new average weighted metric $M'=\sum f_i^* m_i / \sum n_i w_i$, we find that Analyser A has a value of 0.69 and Analyser B a value of 0.78. We can show that this is a statistically significant difference by a t-test given these means assuming a standard deviation of 15% and a sample size of 16 in the survey. The differences between the two analysers is now apparent, particularly when we consider the critical value of 0.75 for minimal performance.

It is acknowledged that there are some limitations with this approach. These include potential differences between performance in EQA and routine performance with patient samples. EQA material is produced to have a range of measurands and therefore some of these may be spiked using non-human or human material which can lead to so-called matrix effects [7]. However matrix effects usually cause differences in method medians which are not used in the metric calculation. Thus most methods produce comparable CVs in EQA schemes to QC CVs, which would be how capability is determined, so the metric is as good as a capability metric.

A second limitation is that for a period of time there will be limited EQA data for new analysers or methods. Thus the benefit of this metric will not be available to early adopters of new instruments but will benefit a significant number of users later in time.

We also acknowledge that the weightings employed in the metric are quite arbitrary and will depend on the individual laboratories needs and priorities. The metric could further be modified to include a cost of test, for example, $M'=\sum(f_i^* m_i+g_i^* c_i)/\sum n_i w_i$, where g_i is a weight per test i (which may be the same as f_i) and c_i is the cost of test i.

Despite the above we believe the metric has some value and suggest that EQA providers could offer these unweighted metrics routinely to provide transparent assay performance indicators. Moreover in the situation where the EQA uses commutable material the metric has no major limitations except for newer analysers where there is a limited history of EQA results. It is important to stress that the metric is not meant to be used exclusively but as another objective tool to be used in what is usually a complex evaluation process.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.

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

References

1. Woodworth A, Korpi-Steiner N, Miller JJ, Rao LV, Yundt-Pacheo J, Kuchipudi L, et al. Utilization of assay performance characteristics to estimate HbA1c result reliability. Clin Chem 2014;60:1073–9.
2. Westgard JO, Westgard SA. Assessing quality on the sigma scale from proficiency testing and external quality assessment surveys. Clin Chem Lab Med 2015;53:1531–5.
3. RCPAQAP Allowable Limits of Performance. https://www.rcpaqap.com.au/wp-content/uploads/2014/11/chempath/docs/Allowable_Limits_of_Performance.pdf. Accessed 5 July 2015.
4. RCPAQAP Calculation of Participant Performance. https://www.rcpaqap.com.au/wp-content/uploads/2013/06/chempath/Notes_on_the_EOCReports.pdf. Accessed 5 July 2015.
5. Fraser CG. Biological variation: from principles to practice. Washington: AACC Press, 2001.
6. RCPAQAP. https://www.rcpaqap.com.au/. Accessed 16 July 2015.
7. Miller WG, Myers GL, Rej R. Why commutability matters. Clin Chem 2006;52:553–4.