Clinical Biochemistry xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com/locate/clinbiochem

Assay Stability, the missing component of the Error Budget

Mark Mackay, Gabe Hegedus, Tony Badrick*

RCPAQAP, St Leonards, Sydney, NSW, Australia

ARTICLE INFO

Keywords:
Error detection
QC rules
Assay bias
Six sigma
Assay Capability
Analytical Performance Specification

ABSTRACT

Gaining a better understanding of Quality Control (QC) processes is a key requirement to improving performance and reducing patient risk. Detecting analytical error is dependent on a QC strategy that reliably detects a critical shift in a result away from the true value.

Recently the concept of Six Sigma has been used by diagnostic laboratories to assess the performance of assays and to assist in the selection of QC rules. The sigma metric is one measure of an assay's ability to perform within specification. However an additional dimension to managing an assay is its stability in bias over time.

The concept of long term stability is the same as measured QC drift (SE_{drift}) which is the effect of numerous calibrations, changes in reagent lots and other conditions i.e. a long term effect. This implies that the standard error budget is wrong because it is modelled on short term QC and misses this SE_{drift} stability component.

We show that SE_{drift} provides a measure of Assay Stability that should be included in Quality Planning and that by including an allowance for this drift, determining target imprecision appropriate for matched QC algorithms that provide high error detection is as simple as dividing the Allowable Performance Specification by 4, 5 or 6.

1. Introduction

The detection of analytical error is dependent on the power of a Quality Control (QC) strategy to detect a critical shift in a result away from the true value. This shift is called the critical Systematic Error (SE_{crit}). The ability to detect SE_{crit} is a function of the actual assay imprecision and bias as well as the QC target imprecision and adopted QC rules relative to the Analytical Performance Specification (APS) or goal. This relative nature is important, i.e. the ratios of imprecision and bias relative to APS are important in the ability to detect error.

The general concept of the capability of an assay is used in many diagnostic laboratories to set QC rules [1,2]. In this paper we use our simple Assay Capability calculation of APS divided by imprecision (Cp_a = APS / SD) which indicates the number of SD inside the allowable limit of performance. Generally a QC protocol involves the QC rules, the QC sample concentration and the frequency of when QC samples are run in a batch [3]. Stability is an extra dimension when managing an assay but is not specifically considered in the standard error budget model used for QC Planning. Stability relates to long term changes, that is changes in bias over months. Stability means that over time the mean of the QC samples does not drift significantly in proportion to the APS. Some assays are very capable (Cp_a > 6) but still exhibit drift and some assay are not very capable but show little drift. Assays that are stable are easier to control and require fewer QC samples, even if they have low capability.

Our aim in this paper is to define long term stability in terms of the components of an error budget [4–6] and the Assay Capability metric Cp_a and explore applications of the concept.

We introduce a new term, SE_{drift}, the error associated with long term calibration and changes in reagent lots and other conditions, i.e. the variability or drift in QC means. The size of SE_{drift} depends on how well assays are managed, i.e. QC drift is larger than it could be if the assays were not well controlled. This implies measuring assay QC drift does not provide a value for acceptable SE_{drift}. We calculate theoretical limits

Abbreviations: APS, Analytical Performance Specification. Previously the term analytical goal was used, but this becomes confused with performance goals as graded levels of performance rather than an allowed limit for error in a result.; Cp_a, Assay Capability defined as the Analytical Performance Specification divided by imprecision, i.e. Cp_a = APS/SD; RCPAQAP, Royal College of Pathologists of Australasia Quality Assurance Programs Pty Ltd; RE_{cont}, Generic term for the random error detected by a Quality Control procedure; SD, standard deviation; SE_{cont}, Generic term for the Systematic Error detected by a Quality Control procedure.; SE_{crit}, Critical Systematic Error. The size of the systematic error that the Quality Control algorithm/procedure must detect to maintain a quality specification, e.g. < 5% of results outside APS.; SE_{drift}, the error associated with long term calibration and changes in reagent lots and other conditions estimated as the variability or drift in QC means. The drift, i.e. systematic error, between estimates of imprecision.; TE_a, total error analytical. An older term for Analytical Performance Specification.; P_{ed}, probability of error detection (i.e. that QC detects an out of control situation); P_{fr}, probability of false rejection (i.e. QC flags a problem but the assay was in control)

* Corresponding author.
E-mail address: tony.badrick@rcpaqap.com.au (T. Badrick).

http://dx.doi.org/10.1016/j.clinbiochem.2017.07.004
Received 19 February 2017; Received in revised form 14 July 2017; Accepted 16 July 2017

for SE_{drift} by examining the effect of drift on error detection, show how well these match actual performance, and use this to simplify QC algorithm selection and improve the QC error budget.

2. Methods

We first studied QC drift when implementing an error budget on a single analyser, and later examined long term retrospective QC data on multiple analysers across a network. We expressed imprecision, bias and drift as fractions of the APS, which was taken as the RCPAQAP Chemical Pathology Allowable Limits of Performance [7]. Data from the long term study was reexamined calculating drift from the overall mean of each QC level.

2.1. Error budget

Error budgets include the sensitivity of the QC procedure to ensure performance at a quality specification within the APS. Typically the specification is that the QC procedure must detect with high probability (90% P_{ed}) and low false rejection rate (P_{fr}) the error or critical shift calculated using stable laboratory measurement bias and imprecision, so that at the critical shift (SE_{crit}), < 5% of results are outside the APS. This turns reactive QC into proactive QC where 95% of results are within APS when the error condition is detected. This reduces patient risk. However, we note that the standard error budget does not specify the timeframe for determining stable laboratory imprecision and bias [5].

2.2. Assay Capability metric, Cp_a

Assay Capability was calculated as Analytical Performance Specification divided by imprecision (Cp_a = APS / CV = APS / SD). This calculation Cp_a is independent of bias or drift, reflects the actual number of standard deviations inside the APS, and converts assay imprecision into a performance variable that can be compared across different QC levels, assays and laboratories. This normalisation technique can also be used to assess performance against different specifications, e.g. CLIA, by simply substituting them for APS.

2.3. Assay Stability metric, SE_{drift}

Bias always includes the sign of the value, i.e. bias of equal size in different directions over time cancels. We consider that drift is different to bias because a reported result with drift is always different from the true value by the amount of drift. In our analysis, the amount of drift is always the size of the shift irrespective of direction.

We calculated SE_{drift} as the difference between QC means and an appropriate target, and expressed it as a percentage of the APS. Our first study of drift used the initial mean as the target. Data from our second study was examined twice; firstly using the QC set mean as the target, and then using the overall mean for the period of the study, which is more statistically appropriate. For the single analyser in common to both our studies, detailed analysis included calculating the average and standard deviation (SD) of SE_{drift} for various combinations of time period, measurand, QC level for each assay and overall.

SE_{drift} estimates drift at one point in time, but what is useful for QC planning is the scatter of these points, which is either presented on a graph or by the standard deviation of SE_{drift}.

2.4. Measuring drift in QC means

We implemented an error budget for 22 common measurands on a single analyser (Hitachi 747) in one laboratory, using short term imprecision to calculate SE_{crit} and Westgard Validator [8] to select appropriate QC algorithms to deliver 90% Ped of the critical shifts at each QC level. After > 20 days, QC means were recalculated and used as the target for assessing subsequent drift in each QC mean (SE_{drift}) over sequential time periods of 16, 15, 25, 49 and 6 days. Drift was expressed as a percentage of the APS for that QC level.

To confirm the findings of the initial study, a longer term study of Assay Stability examined retrospective monthly QC data from four different analyers each in a separate laboratory in a network of one of the authors [9]. One analyser was the same instrument from the first study, with QC data starting 12 months after the end of the first study and extending for 14 months. We examined 6 or 7 months of QC data from the other three instruments, which used dry slide technology (V250 or V950 from Ortho Clinical Diagnostics). Drift was calculated as the difference between the analyser set target mean and the QC means at each data accumulation period expressed as a percentage of the APS.

On the analyser in common with the first study (Hitachi 747), 17 of the 22 assays in the network study had QC means reset sometime after 9 months. To remove the effect of resetting QC targets, this data was reanalysed using data only for the period up until the QC mean was reset (being 9–14 months depending on the assay) and drift was calculated as the difference between the monthly mean and overall mean for that QC level.

For all studies, the QC data collected represented 'in control' data, i.e. standard laboratory practice was to eliminate any failed run QC data before data accumulation.

2.5. Accuracy of QC mean

The reliability of each QC mean affects the accuracy of each SE_{drift} measurement. The Standard Error of the Mean (SEM) is a measure of closeness of a calculated mean from a particular data sample to the true mean calculated from the whole data population and is calculated as $SD/\sqrt{n}$. The lower the SEM, the more accurate SE_{drift}. As imprecision declines, SD increases and SEM increases, which is the expected behaviour of SE_{drift}. We examined the effect of SEM on SE_{drift} by calculating the ratio of SE_{drift} to SEM to show their relative sizes. Results were plotted against SE_{drift}, with points grouped into three Assay Capability imprecision classes to show the effect of increasing imprecision.

2.6. External quality assurance

If QC and EQA estimates of imprecision can be aligned, data on peer imprecision can be used for QC Planning. RCPAQAP Chemical Pathology end of cycle imprecision calculations use data collected over months, but sample size is small compared to QC. We examined the adjustment required to accommodate the small sample size of EQA estimates of long term imprecision, and compared it to SE_{drift} which in effect converts QC short term imprecision into long term imprecision.

2.7. Relationships between Assay Capability, Assay Stability and error detection

We used a standard error budget to derive the simple relationship between the critical shift the QC procedure must detect, assay capability and Assay Stability, i.e. between SE_{crit}, Cp_a and SE_{drift}. We examined the effect of SE_{drift} on error detection (P_{ed}) at different levels of imprecision by calculating SE_{crit} at specific combinations of Cp_a and SE_{drift} for common QC algorithms and then estimating P_{ed} from Westgard critical error graphs. Data described the performance boundaries for SE_{drift}.

3. Results

The extent of QC drift immediately after implementing an error budget is shown by displaying the distribution of SE_{drift} versus Cp_a for all QC assay levels. The scatter pattern does not appear to be altered by measurand concentration (Fig. 1) or the time period of data accumulation (Fig. 2).