Clinica Chimica Acta 486 (2018) 329–334

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

Quality control limits: Are we setting them too wide?

Robert L. Schmidt[a], Brandon S. Walker[b], Lauren N. Pearson[a,*]

[a] Department of Pathology, ARUP Laboratories, University of Utah, Salt Lake City, UT, United States
[b] ARUP Laboratories, Salt Lake City, UT, United States

ARTICLE INFO

Keywords:
Statistical Quality Control
Variability
Process Improvement
QC
Variation
QC limits

ABSTRACT

Background: Quality control charts (Levey Jennings Charts) are based on estimates of variation. There are two general approaches for estimating variation: those based on short-term variation and those based on long-term variation. We have observed that clinical laboratory science (CLS) tends to estimate variation using long-term variation but that most other fields use short-term variation. The objective of this study is to compare these two methods of estimating process variation, compare the accuracy of control limits generated by each method, and explore whether it would be useful for clinical laboratories to adopt methods used in other fields.

Methods: We conducted a literature review to compare recommendations for methods for estimation of variation in CLS with other fields. We searched textbooks for suggested methods and also searched the primary literature for references to methods associated with short-term and long-term variation. We provide theoretical results from statistics to show that, in theory, short-term estimates can differ from long-term estimates of variation. We used simulation studies to show that one can construct examples where short-term and long-term estimates of variation lead to significant differences in control limits. Finally, we show laboratory data comparing short-term and long-term estimates of variation.

Results: We found that practice in CLS differs from other fields. We found no references to methods based on short-term variation in CLS textbooks and only one reference in the primary literature. In contrast, standard quality control (QC) texts recommend methods based on short-term variation and the primary literature makes frequent reference to such methods. We found statistical papers that show that, in theory, estimates based on long-term variation can produce inflated estimates of process variation. We used simulation to show that such examples can be constructed. We examined 95 QC charts and found that in 93 cases, there were significant differences between short-term and long-term estimates of variation. The ratio of long-term to short-term variation was greater than 1.5 in 18% of cases.

Conclusion: Estimates of variation based on short-term and long-term variation can lead to significant differences in estimates. Estimates based on long-term variation are frequently larger than estimates based on short-term variation.

1. Introduction

Quality control (QC) is a critical activity because it is used to ensure the reliability of patient results. Laboratories generally use control charts to assess the reliability of results. Control charts are used to assess process stability. A stable process will have variation, but the variation is consistent and predictable. This predictable variation is used to establish control limits. Control limits play a central role in quality control because they are used to define normal vs abnormal variation and, for that reason, it is important that control limits are accurate.

Control limits are based on estimates of process variation. Thus, it is important to use the best methods to estimate process variation. We have observed that the methods clinical laboratories use to estimate variation differ from those commonly used in other fields. Specifically, clinical laboratories generally estimate process variation directly from a runs chart (Levy-Jennings or LJ chart). These estimates are a measure of long-term variation. Outside of clinical laboratories, estimates of variation are based on measures of short-term variation. Estimates of short-term variation are generally obtained from range charts (R charts) or s charts. In this paper, we compare these two methods of estimating process variation, compare the accuracy of control limits generated by each method, and explore whether it would be useful for clinical laboratories to adopt methods used in other fields.

* Corresponding author at: Department of Pathology, University of Utah, 15 N Medical Drive East, Suite 1100, Salt Lake City UT 84112, United States.
E-mail address: lauren.pearson@aruplab.com (L.N. Pearson).

https://doi.org/10.1016/j.cca.2018.08.031
Received 11 June 2018; Received in revised form 17 August 2018; Accepted 20 August 2018
Available online 23 August 2018
0009-8981/

2. Methods

2.1. Theoretical background

We provide background material on the analysis of process variation and the use of control charts. We provide this background because the general quality control literature uses terms and concepts that are not generally used in the clinical laboratory QC literature. We refer to these different practice patterns as clinical laboratory QC (CLQC) and general QC (GQC).

2.2. Literature search

To examine differences in QC practice, we searched for references to R charts in standard texts on CLQC as well as standard texts in the GQC literature. We searched for R charts because R charts are generally used to estimate short-term variation. We hypothesized that references to R charts would be associated with different patterns of practice: fields using short-term estimates of variation would make references to R charts whereas those that use long-term estimates would not make frequent references to R charts. We also searched Scopus (an electronic database that spans PubMed and other databases) for references to "R-chart", "R chart" or "Range Chart" in the primary research literature. We counted the total number of references to R charts as well as the number of references that appeared in journals that focus on clinical laboratory science, analytical chemistry, or medicine. All searches were performed on July 26, 2018.

2.3. Simulation studies

We used simulation to demonstrate that estimates of the standard deviation based on short-term and long-term variation can differ. To that end, we created a simulated dataset with 1000 observations drawn from a normal distribution (mean = 0, standard deviation = 1.0). This underlying variation was used to represent the intrinsic variation of a well-controlled process. We added shifts of 2 standard deviations to represent out-of-control variation. We then used the resulting LJ chart and R chart to generate estimates of the standard deviation.

2.4. Laboratory data

We used actual QC data to demonstrate that estimates of variation (and control limits) based on LJ charts differ from estimates based R charts. To this end, we examined QC data for 95 combinations of analytes and QC levels. All of these assays are performed by mass spectrometry in one of our toxicology laboratories and represents all of the assays performed in that laboratory. We used the SR test to determine whether there was a statistically significant difference between the estimates [1]. The SR test evaluates the ratio of the standard deviation estimated by the LJ-chart to the standard deviation estimated by the R chart.

3. Results

3.1. Theoretical background

Control charts are based on concepts of variation. The output of a stable process has variation but the variation is predictable and occurs within certain limits. This type of variation is known as common cause variation [2–4]. Common cause variation reflects the effect of many small sources of variation in the measurement process which, when combined, give rise to variation in the final measurement (Fig. 1). It is impossible to link common cause variation to any particular cause because it is the result of many small contributions. Common cause variation contains no information or patterns [5]. Common cause variation reflects the natural variation that occurs when the process operates as designed [6].

Control charts are designed to detect process instability. Results may be unreliable when a process is unstable. Instability is detected by observing departures from common cause variation. These departures are due to so-called assignable or special cause variation (Fig. 2). In principle, assignable cause variation can be attributed to a change in an input that deviates from the normal operation of the process (e.g., an equipment malfunction, failure to follow a procedure, etc.) The presence of assignable cause variation implies that the measurement process is not operating as designed and that some extraneous factor is acting on the process. Results may be unreliable when assignable cause variation is present and, consequently, it is important to have systems in place to detect assignable cause variation. By definition, a process is unstable if assignable cause variation is present [4].

A process can have many sources of variation. Assignable cause variation is related to sources of variation that have a measurable impact on the final result. In principle, these sources of variation can be identified and controlled. Process improvement is based on learning about and controlling assignable cause variation [2,3,7,8].

Control charts are designed to monitor a process and detect the presence of abnormal or assignable cause variation. Normal variation is defined by control limits. For example, the chance of a stable process producing a result that exceeds three standard deviations is very unlikely (3 in 1000). Thus, when such a result is observed, it is reasonable to question whether the result was produced by a stable process. This is the origin of the so-called 1–3 s rule: the process is considered unstable if one QC result exceeds three standard deviations. All control chart methods are based on similar logic. They look for deviations from stable, common cause variation. A process is said to be in statistical control when no such deviations (i.e., assignable cause variation) are present.

Control limits are based on estimates of "normal" process variation or common cause variation. Data containing assignable cause variation will inflate the estimate of the standard deviation leading to inappropriately wide control limits [2,9–20]. Thus, the standard deviation should be estimated when a process is in control. This is challenging because stability is defined in terms of control limits which, in turn, are based on the data that is being evaluated for stability [9,21].

The sample size required to provide accurate estimates of the standard deviation poses additional challenges. Clinical & Laboratory Standards Institute suggests that a sample size of 20 is sufficient, but research suggests sample sizes of at least 100 are required to obtain accurate parameter estimates [21–24]. This is challenging because larger sample sizes provide greater opportunities for data to incorporate assignable cause variation [21]. Given these challenges, how should the standard deviation be estimated?

There are two main methods for estimating the standard deviation of QC values on an LJ chart [9]. The first method assumes that the data originate from a homogeneous group and calculates the sample standard deviation using the entire data set:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}} \quad (1)$$

The estimate of the population standard deviation is given by $\hat{\sigma}_L = s/c_4$ where c_4 is a constant chosen to make the estimator unbiased. This can be considered a long-term estimate of the variation. This variation contains two components: common cause and assignable cause variation.

The second method makes use of rational subgrouping and combines estimates of the standard deviation from each subgroup. The estimates are based on the average range of the subgroups.

$$\hat{\sigma}_s = \overline{R}/d_2 \quad (2)$$

where d_2 is a constant that depends on the sample size. This estimate can be considered an estimate of short-term variation. Most texts on