Focus Article:
How to establish the mean and standard deviation of internal quality control samples to construct control charts.
The Bayesian approach with an example of D-dimer

Panagiotis Tsiamyrtzis[1], Frédéric Sobas[2] and Claude Négrier[2]

1: Dept. of Statistics, Athens University of Economics and Business, Athens, Greece

2: Hemostasis Laboratory, Edouard Herriot Hospital, Lyon, France

Abbreviations:
FAP: False alarm probability
IC: In control
IID: Independent identically distributed
IQC: Internal quality control
BCC: Bayesian control chart
SPC: Statistical process control
$\hat{\mu}$: Estimated prior target mean value for the control material under monitoring
$\hat{\sigma}$: Estimated maximum acceptable inter-assay standard de viation (SD) according to the manufacturer
$\hat{\tau}$: Estimated inter-assay SD during method validation in each laboratory

Introduction

Internal quality control (IQC) is an important tool for establishing whether an analytical test system produces reliable results. For IQC one or more quality control (QC) samples are run prior to or simultaneously with patient samples. If the results of the control samples are within certain predefined acceptance limits, the results of patients' samples can be released. QC sample test results are frequently displayed in so-called control charts. To evaluate these QC sample results and look into trends statistical process control (SPC) methods are used. One of the well-known QC rules is the 1_{2s} QC rule [1,2] which signals anything that exceeds two SD from the expected value. For each new batch of IQC samples the expected value (mean value) and SD need to be established to construct a control chart. This requires two phases: I (preliminary) & II (testing). The former is done in an off-line mode (i.e. the QC sample is not yet used for making a decision about the status of an analytical system), where first we estimate the parameters used to build the chart and then examine the data retrospectively, i.e. once phase I is completed all the phase I data will be examined for conformance with the established limits. On the other hand phase II runs on-line, i.e. each new reading is plotted on the chart and on-line inference is available. It is important not to forget the statistical aspect of the IQC management regarding the assumption of approximating the normal distribution when enough data are observed. In other words the phase I management is very crucial with respect to the conventional phase II reliability.

In this contribution we focus our attention on the conventional preliminary phase, where the major goals are both to be able to perform efficient QC monitoring even when we have very few data points available and to obtain "reliable" estimates of the mean and the inter-assay SD for the next long-term conventional QC management. It is possible to run an off-line preliminary phase in advance, using new control batches before actually changing batches. But there is considerable technical and economic interest in getting round this conventional preliminary phase management with regard to the high number of laboratory tests especially when the measurement series are not frequent. Unreliable estimates of the mean and the inter assay SD can have from serious to catastrophic results on the performance of the classical control chart in both phases I & II [3,4]. This means that QC results can either be falsely rejected or accepted. This implies the risk of a waste of resources (falsely rejected) or the risk of a wrong interpretation of patient sample results (falsely accepted).

The classical approach during the preliminary phase (off-line method), assumes the process is in the in-control (IC) state (i.e. no abnormal cases should be present), with independent identically distributed (IID) observations. The longer the preliminary phase, the more accurate the estimates, but simultaneously the more likely it is that the process will deviate from its IC state. Typically, the mean and inter-assay SD are estimated from at least 20 and 30 control values respectively [5]. The preliminary phase uses up a great deal of resources, given the large panel of tests a laboratory has to carry out. This is especially true as laboratory examinations are rarely performed in continuous series, resulting in overlaps with currently used QC samples that are very hard to manage.

This work proposes an alternative monitoring mechanism that will not require a preliminary phase, allowing on-line inference from the beginning (i.e. from the second measurement onwards). It is based on the Bayesian logic where we utilize available manufacturer's prior information. The two main manufacturer prior parameters are:

1. Prior target mean of assayed quality control materials (manufacturer control materials with assigned values).
2. Maximum acceptable inter-assay SD value on methods with reagents and device both provided by the manufacturer (technical notices specifying the maximum acceptable inter assay SD)

We provide an Excel spreadsheet (downloadable on the ECAT website) where the proposed monitoring method can be applied to the process readings once we specify three parameter values:

i) $\hat{\mu}$: the prior target mean value for the assayed control-material being monitored (i.e. the target value provided by the manufacturer).

ii) $\hat{\sigma}$: the maximum acceptable SD specified by the manufacturer.
iii) $\hat{\tau}$: each laboratory's inter-assay SD determined during the method-validation phase (with at least 30 data points [5]).

The data can be immediately used for on-line monitoring. The data is put sequentially in the Excel file, as they arrive and have immediate inference for the process, (i.e. "ALARM" if there is a loss of statistical control state). At the end of phase I the Excel tool provides a mean (when we have at least 20 observations in phase I), which can then be implemented in a conventional control chart for the testing phase. With 30 data points, the Excel tool will also calculate the inter-assay SD value that one can compare to the inter-assay SD determined during the validation phase of the method in order to identify an underlying matrix effect [6].

Methods

Estimating the required parameter values for the Bayesian control chart

The construction of the Bayesian control chart (BCC) requires three parameters: $\hat{\mu}$ and $\hat{\sigma}$ which are obtained from manufacturer specifications and $\hat{\tau}$ which reflects the accuracy of measurements in the laboratory and is established during the laboratory validation process.

For estimating $\hat{\mu}$ and $\hat{\sigma}$, we can use the fact that the manufacturer normally provides the acceptance range for the IQC results (i.e. the mean +/- 2 * SD) and a coefficient of maximum acceptable variation CV for a given process. The data have a high probability of being within these limits if the process is under the in-control state. Then an estimate of $\hat{\mu}$ and $\hat{\sigma}$ can be calculated by :

$$\hat{\mu} = \frac{L+U}{2} \quad \text{and} \quad \hat{\sigma} = \hat{\mu} \times CV$$

(L = lower limit of acceptance range; U = upper limit of acceptance range)
Example : L = 0.8 IU/mL ; U = 1.0 IU/mL ; CV = 5 %
$\hat{\mu}$ = (0.8 + 1) / 2 = 0.9 ; $\hat{\sigma}$ = 0.9 * 0.05 = 0.045
The parameters are used in the Excel template.

Regarding $\hat{\tau}$, i.e. the accuracy of the measurement, it depends on various laboratory factors: equipment, experience of the technician, etc. It can be assessed during in-laboratory method validation upstream of implementation, as the inter-assay SD, varying according to the degree of control over these nuisance parameters [7].

As data are obtained sequentially, they are entered into the Bayesian chart and from the second observation it is valid to perform inference. During the preliminary phase the parameters $\hat{\mu}$ and $\hat{\sigma}$ are continuously updated based on the actual IQC results obtained. The initial prior settings of $\hat{\mu}$ and $\hat{\sigma}$, will affect the performance of the chart only for the very first few observations, and their effect will vanish as more data become available. Thus, as long as we avoid extreme choices (such as very small $\hat{\sigma}$), the chart will be quite robust even when poor estimates were used for these prior settings. On the other hand, the chart will be more sensitive to the parameter $\hat{\tau}$, which reflects the accuracy of the laboratory, since it will be fixed and not updated at any stage. A sensitivity analysis in the Results section examines the effect of parameter misspecification in the case study considered.

Technical details of the BCC construction, along with an Excel BCC template that runs the suggested methodology can be found in the member section of the ECAT website and can be tested by interested users.

Results

Case study: Reagent and automated coagulation analyzer

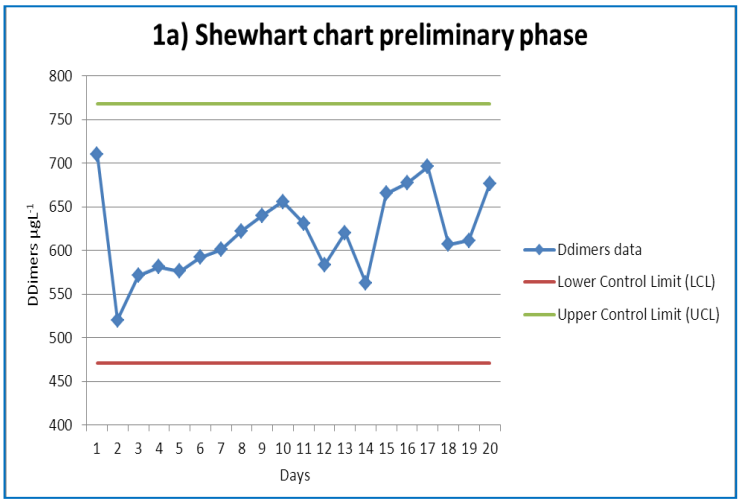
An Instrumentation Laboratory (IL) (Bedford, MA, USA) automate and reagents were used (analyzer: ACLTOP 500 CTS®; reagent: D-Dimer HS 500® for D-Dimer quantification in citrated human plasma). Control material was a low-control sample (D-Dimer HS 500® control level 1), with prior allowable inter-assay SD defined by IL as 65 ($\hat{\sigma}$) and prior mean value 544 μgL^{-1}.

Acceptability of the 20 control values was confirmed as they had been collected during a phase of overlap with another control batch with the same reference, in turn collected respecting the 1_{3s} rule. In the new IQC batch, IL's prior target value was 544 μgL^{-1}, and the maximum acceptable inter-assay SD should not exceed 65 when validated in each laboratory. In the checking phase, inter-assay SD in our laboratory was 49 ($\hat{\tau}$) [7].

Analysis of preliminary phase data

After the preliminary phase, performed on the 20 values, the mean was 620 μgL^{-1} and inter-assay SD 49, while the control limits were set at ±3.016 so that a 5% false alarm probability (FAP) was achieved for the whole sequence of 20 values (Figure 1a). This is called the 1_{3s} method. To be equivalent to

Figure 1. a) Shewhart chart and b) Bayesian control chart (BCC) during preliminary phase



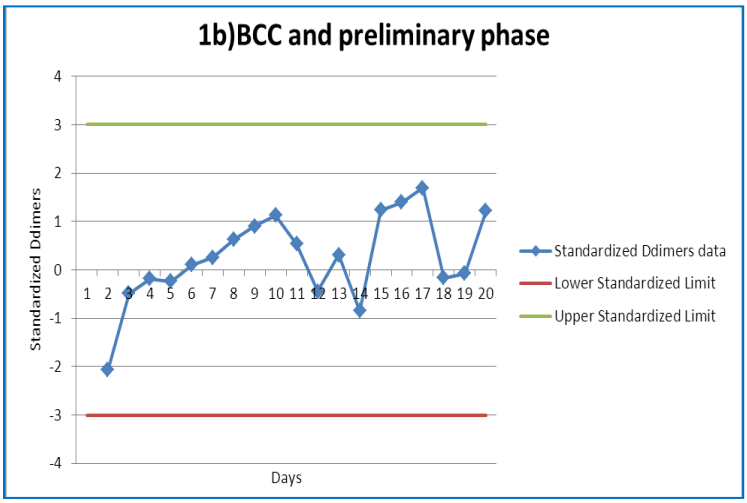
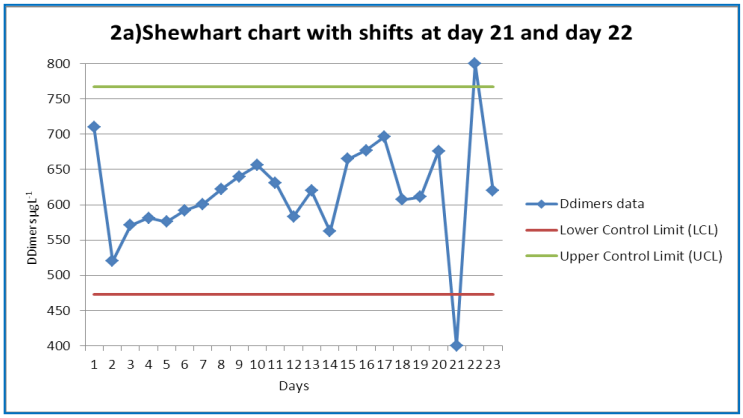


Figure 2. a) Shewhart chart and b) Bayesian control chart (BCC) with shifts at day 21 and day 22



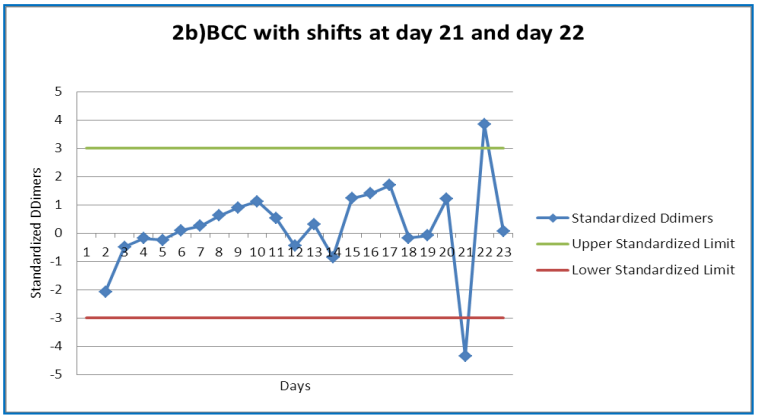
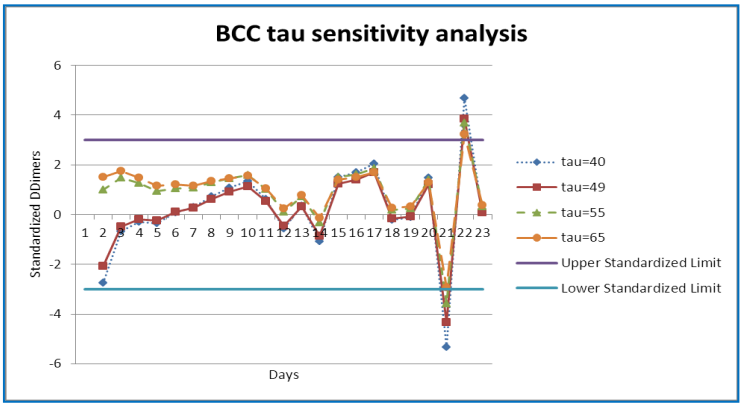


Figure 3. Bayesian control chart (BCC) with shifts at day 21 and day 22 and sensitivity analysis on « own inter assay SD estimated at method validation phase (tau) »



the BCC, the 1_{3s} chart was selected which plots the control limits at 3.01599 SD from the center line, achieving a 5% overall FAP.

The BCC identified no outliers during a preliminary phase well controlled by the overlap phase: i.e., the BCC was not subject to false rejection with $\hat{\tau}$ =49 (Figure 1b).

Analysis of data after the preliminary phase with a shift scenario

The 1_{3s} rule and BCC with $\hat{\tau}$ =49 detected results outside the established limits simulated after the preliminary phase (Figures 2a and 2b).

Examining the sensitivity to the parameter estimates we found that the BCC detected simulated shifts after the preliminary phase with $\hat{\tau} < 65$, while for $\hat{\tau} > 65$, it no longer detected all alarms (Figure 3). This is expected as the large value of $\hat{\tau}$ is associated with rather inaccurate laboratory measurements that can help outlying observations to escape detection. For the remaining two parameters, $\hat{\mu}$ and $\hat{\sigma}$ the sensitivity analysis showed very minor differences as we alter them. Furthermore, these differences were observed only in the very first few data in the process.

Discussion

With the introduction of a new batch of control samples it is necessary to run a preliminary phase. During this phase QC data for accepting or rejecting an analytical was run with the ongoing batch. However, running a preliminary phase can be costly and time-consuming, especially when particular measurements are not frequently performed [4]. Therefore there is considerable technical and economic interest in getting round this problem.

The laboratory may be thought to focus only on manufacturer specifications to define control value acceptability. However it is not unreasonable to use manufacturer specifications (i.e. manufacturers' prior target values and allowable analytic performance are derived from plentiful data from multiple machines and batches), if the analytic system is a good one.

Laboratories with good analytic practice (small inter-assay SD) benefit most from BCCs (Figure 3). They are better able to detect outliers than laboratories with poorer analytic performance. Laboratories must therefore be as careful as possible in estimating prior inter-assay SD in the validation phase.

The Bayesian approach is also useful for both methods which are rarely done and methods using a small batch of IQC samples. It is possible to monitor this kind of method with this short-term Bayesian approach.

Thus, both theoretically and practically, the laboratory is bringing its method under control as soon as it begins implementing its IQC values. At least this Bayesian model can serve as a complement to a conventional approach, which can be reintroduced as soon as there are enough reliable IQC data.

Acknowledgements

The authors would like to thank Dr P. Meijer (ECAT Foundation, Leiden, Netherlands), Dr W.L. Nichols (Mayo Clinic College of Medicine, Rochester, MN, USA) for their kind support and Iain Mc Gill for the translation.

References

1. Housley D, Kearney E, English E, Smith N, Teal T. Audit of internal quality control practice and processes in the south east of England and suggested regional standards. Ann Clin Biochem 2008; 45:135-9
2. Parvin CA, Kuchipudi L, Yundt-Pacheco JC. Should I repeat my 1:2s QC rejection? Clin Chem 2012; 58:925-9
3. Montgomery CD. Introduction to Statistical Quality Control 7th Ed New York: Wiley 2012
4. Marquis P. Internal quality control: false rejects and preliminary phase. Ann Biol Clin 2001; 59:214-18
5. C24A3. Statistical quality control for quantitative measurement procedures: principles and definitions; approved guideline - third edition CLSI 2006
6. Miller WG, Erek 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
7. EP05-A2. Evaluation of precision performance of quantitative measurement methods; approved guideline-second edition CLSI 2004