

---

## Assessing the Impact of the Frequency of Quality Control Testing on the Quality of Reported Patient Results

Curtis A. Parvin<sup>1</sup>

---

**BACKGROUND:** The traditional measure used to evaluate QC performance is the probability of rejecting an analytical run that contains a critical out-of-control error condition. The probability of rejecting an analytical run, however, is not affected by changes in QC-testing frequency. A different performance measure is necessary to assess the impact of the frequency of QC testing.

**METHODS:** I used a statistical model to define in-control and out-of-control processes, laboratory testing modes, and quality control strategies.

**RESULTS:** The expected increase in the number of unacceptable patient results reported during the presence of an undetected out-of-control error condition is a performance measure that is affected by changes in QC-testing frequency. I derived this measure for different out-of-control error conditions and laboratory testing modes and showed that a worst-case expected increase in the number of unacceptable patient results reported can be estimated. The laboratory thus has the ability to design QC strategies that limit the expected number of unacceptable patient results reported.

**CONCLUSIONS:** To assess the impact of the frequency of QC testing on QC performance, it is necessary to move beyond thinking in terms of the probability of accepting or rejecting analytical runs. A performance measure based on the expected increase in the number of unacceptable patient results reported has the dual advantage of objectively assessing the impact of changes in QC-testing frequency and putting focus on the quality of reported patient results rather than the quality of laboratory batches.

© 2008 American Association for Clinical Chemistry

---

One of the big issues confronting today's clinical laboratories is the desire to objectively address the question

of how frequently QC testing should be performed (1). The traditional approach to assessing quality control performance is based on the probability of rejecting an out-of-control analytical run, where an analytical run is defined as the group of patient specimens for which a decision about control status is being made (2). The probability of rejecting an analytical run depends on the QC rule that is applied, the number of QC sample results used by the rule, and the true state of the testing process—whether the process is in control or the degree to which it is out of control. The probability of rejecting an analytical run, however, does not depend on the frequency of QC testing. It only answers the question, Given that a QC rule is evaluated, what is the probability of a QC rejection?

To objectively address questions related to the frequency of QC testing, an alternative performance measure is needed. Ideally, the performance measure should focus on the quality of individual patient results and should be affected by changes in the frequency of QC testing. A number of possible performance measures could meet the dual requirements of being influenced by changes in the frequency of QC testing and reflecting the quality of reported patients' results. The purpose of this article is to define and evaluate one such quality control performance measure: the expected increase in the number of unacceptable patient results reported during the existence of an undetected out-of-control error condition. I derived the performance measure under different laboratory testing modes and assuming different characteristics for the out-of-control error condition. I then used the performance measure to assess the impact of changes in the frequency of QC testing in addition to changes in the number of QC samples tested and the QC rules applied. Finally, I showed how the performance measure can be used to design optimal QC strategies.

---

<sup>1</sup> Washington University School of Medicine, Department of Pathology and Immunology, Saint Louis, MO.

Address correspondence to the author at: Washington University School of Medicine, Department of Pathology and Immunology, Saint Louis, MO 63110. Fax

(314) 454-5208; e-mail parvin@wustl.edu.

Received July 1, 2008; accepted September 19, 2008.

Previously published online at DOI: 10.1373/clinchem.2008.113639

---

## Materials and Methods

The findings presented here are based on a statistical model of the analytical testing process. The statistical model defines the in-control process, out-of-control processes, laboratory testing mode, and the quality control strategy.

A measurement procedure's in-control state is modeled by a gaussian distribution of measurement errors around the true concentration of a sample. It is assumed that the stable analytical imprecision of the measurement procedure is known at each QC concentration level; the analytical SD at other concentrations is obtained by linear interpolation. Out-of-control error conditions are traditionally modeled as either a shift in the distribution of measurement errors [systematic error (SE)<sup>2</sup>] or an increase in the stable analytical imprecision of the measurement procedure [increase in random error (RE)]. In this article, only systematic error conditions are illustrated, but the methodology is equally applicable to out-of-control error conditions defined by increases in analytical imprecision.

I investigated 2 different laboratory testing modes: batch-mode and continuous-mode processes. For a batch-mode process, a batch (analytical run) includes both patient specimens and QC samples. The key concept from a quality control perspective is that the process is operating in *batch units*. A batch is either in control or out of control. QC rules applied to the QC samples in each batch determine whether to accept or reject a batch. For a continuous-mode process, patient specimens are tested in a continuous stream. Periodically, QC events occur. A QC event is defined as the point in the testing stream at which QC samples are measured and a QC decision is made. The key quality control concept for a continuous-mode process is that the process is operating in *sample units*, and an out-of-control error condition might occur at any point in the testing stream.

I investigated 2 types of out-of-control error conditions. For batch-mode processes, I considered both intermittent and persistent out-of-control errors (3). An intermittent error affects only a single batch. Even if the QC rules fail to detect it, the intermittent out-of-control error condition does not carry over to the next batch. On the other hand, a persistent out-of-control error continues to affect subsequent batches until the error condition is detected. For continuous-mode processes, I considered only persistent error conditions.

For the QC rules investigated here, I mathematically computed the probabilities of QC rule rejection using normal and  $\chi^2$  distribution functions (4). All computations were performed using the Matlab software package (The MathWorks Inc.).

The quality of a laboratory result is related to the magnitude of the difference between the true concentration of a sample and the value reported by the laboratory:  $E = X - T = B + SE + RE \cdot e$ , where  $X$  is the reported concentration,  $T$  is the true concentration,  $B$  is any inherent laboratory bias,  $SE$  is any undetected systematic error,  $RE$  is any undetected increase in stable analytical imprecision, and  $e$  is gaussian-distributed random measurement error reflecting the analytical SD,  $\sigma_a$ , of the measurement procedure. In general,  $B$ ,  $SE$ ,  $RE$ , and  $\sigma_a$  can all vary with concentration. Often, it is assumed that  $B$ ,  $SE$ , and  $RE$  are constant multiples of  $\sigma_a$ .

An unacceptable patient result is defined using the concept of total allowable error (5). Define total allowable error ( $TE_a$ ) as the magnitude of error that is considered unacceptable. The increase in the probability of producing an unacceptable patient result ( $\Delta P_E$ ) is computed as the probability that a result's error exceeds  $TE_a$  in the presence of an out-of-control error condition minus the probability that a result's error exceeds  $TE_a$  when the process is in control. Computation of  $\Delta P_E$  is described in more detail in an earlier report (6).

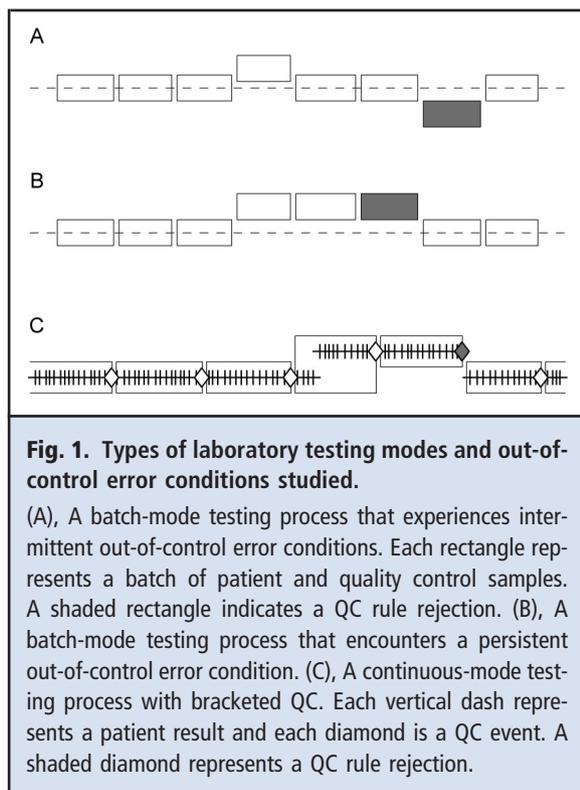
The following example of a laboratory test procedure is used throughout this article. There are 2 concentration levels of control material. One QC level has a concentration of 43 g/L with analytical SD 1.3 g/L and CV 3.0%. The other QC level has a concentration of 80 g/L with analytical SD 1.8 g/L (CV 2.25%). No inherent laboratory bias exists ( $B = 0$ ). The total allowable error for the test procedure is specified to be  $\pm 10\%$  of the true concentration in a sample.

## Results

Laboratory testing modes and out-of-control error conditions investigated are illustrated in Fig. 1. Fig. 1A depicts a batch testing process that experiences intermittent out-of-control error conditions. The horizontal dashed line denotes the in-control state. Each rectangle represents a batch of patient and control samples (analytical run). A shaded rectangle implies that the QC rules rejected the batch. The fourth batch is out of control, but the QC rules fail to detect it. Because of the intermittent nature of the out-of-control error condition, however, the error does not carry over to the next batch (even though it is not detected). Fig. 1B depicts a batch process with a persistent out-of-control error condition. Again, an out-of-control error occurs at the fourth batch, but the QC rules fail to detect it. In this

---

<sup>2</sup> Nonstandard abbreviations: SE, systematic error; RE, random error;  $E(N_U)$ , undetected out-of-control error condition.



case, the error persists in subsequent batches until a QC rejection occurs. Fig. 1C depicts a continuous testing process using a bracketed QC strategy. With bracketed QC, patients' results measured after a passing QC event are held until the next QC event. Only if the subsequent QC event passes are the patients' results reported. In Fig. 1C, each vertical line represents a patient result and each diamond marks a QC event, with a shaded diamond denoting a QC rule rejection. The rectangles define the bracketed patient results that are held until the subsequent QC event passes. An out-of-control error condition occurs within the fourth bracket, but the subsequent QC event fails to detect it. Consequently, the bracketed patient results (some of which are affected by the error condition) are reported. The persistent nature of the error condition means that it continues to affect subsequent patient results. When a QC event finally detects the error condition, those bracketed patient results associated with the failed QC event are not reported.

The expected increase in the number of unacceptable patient results reported during the existence of an undetected out-of-control error condition,  $E(N_U)$ , depends on the laboratory testing mode, the total allowable error specification, the distribution of patient results encountered by the laboratory, the magnitude of an out-of-control error condition, the number of QC

samples measured at each QC event, the QC rules applied to the QC sample results, and the length of the interval between QC events (QC frequency). For a batch-mode process that experiences an intermittent out-of-control error condition:

$$E(N_U) = \Delta P_E(1 - P_{ED})E(N_B).$$

In this equation,  $\Delta P_E$  represents the increase in the probability of producing unacceptable patient results because of the presence of the out-of-control error condition,  $P_{ED}$  the probability of a QC rule rejection, and  $E(N_B)$  the expected number of patient specimens tested in the analytical run. For a batch-mode process with a persistent out-of-control error condition:

$$E(N_U) = \Delta P_E(ARL_{ED} - 1)E(N_B).$$

$ARL_{ED}$  is the average run length to error detection (the average number of batches that are processed before a QC rule rejection occurs).

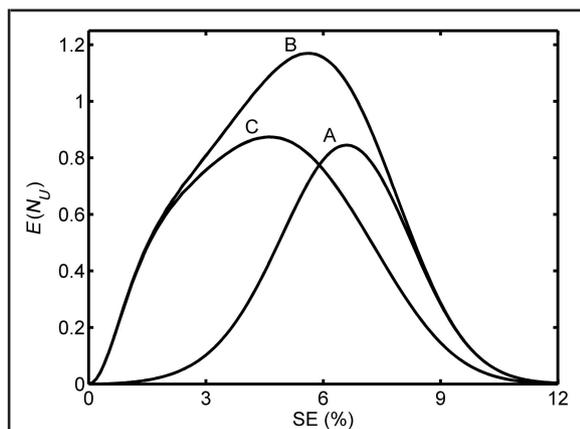
For a continuous-mode process using bracketed QC that encounters a persistent out-of-control error condition:

$$E(N_U) = \Delta P_E \{ (ARL_{ED} - 1)E(N_B) - (1 - P_1)[E(N_B) - E(N_0)] \}.$$

$E(N_0)$  is the expected number of patient results produced between the time an out-of-control error condition occurs and the next QC event, and  $P_1$  is the probability of a QC rule rejection at the first QC event after the out-of-control error condition occurs.

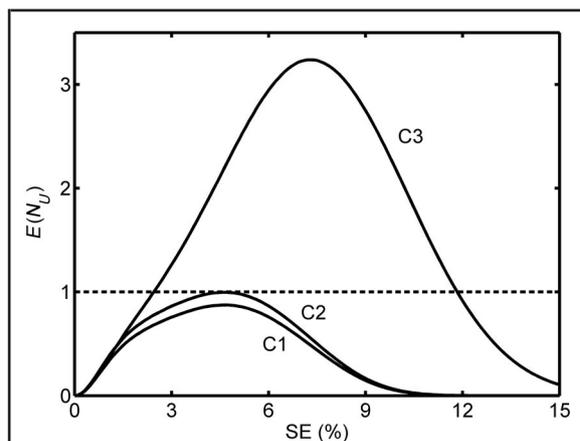
In the above equations,  $\Delta P_E$  is a function of the magnitude of the out-of-control error condition and the total allowable error specification;  $P_{ED}$ ,  $P_1$ , and  $ARL_{ED}$  depend on the magnitude of the out-of-control error condition and the power of the QC rule; and  $E(N_0)$  depends on the interval length between QC events. More detail regarding the derivations of the above equations is given in the Appendix in the Data Supplement that accompanies the online version of this article at <http://www.clinchem.org/content/vol54/issue12>.

Fig. 2 plots  $E(N_U)$  as a function of the magnitude of a systematic out-of-control error condition for the 3 cases. Two QC samples are measured in each batch (or each QC event) and a  $\bar{Z}_{1.98}/S_{2.81}$  QC rule is applied to the 2 QC sample results. This QC rule tests the mean and sample SD of the standardized QC results and is equivalent to a mean/range rule when testing 2 QC samples but gives slightly superior performance with  $>2$  QC samples. The QC rule is more fully described in the Appendix in the online Data Supplement. The false-rejection probability ( $P_{fr}$ ) for this rule is 0.01. In all 3 cases,  $E(N_U)$  starts at 0 when the process is in control, increases as the magnitude of the out-of-control error condition increases, reaches a maximum value, and then decreases



**Fig. 2.** The expected increase in the number of unacceptable patient results reported,  $E(N_U)$ , as a function of the magnitude of a systematic out-of-control error condition, SE, given as a percentage of the true concentration.

Curves are labeled to match the corresponding cases in Fig. 1. In all cases a  $\bar{Z}_{1.98}/S_{2.81}$  QC rule is evaluated with 2 QC samples per batch (per QC event) and 50 patient specimens per batch (between QC events).



**Fig. 3.** The expected increase in the number of unacceptable patient results reported,  $E(N_U)$ , as a function of the magnitude of a systematic out-of-control error condition for different QC utilization rates assuming a continuous-mode testing process with bracketed QC.

Curves C1 and C2 reflect a  $\bar{Z}_{1.98}/S_{2.81}$  QC rule with 2 QC samples per QC event. For curve C1, 50 patient specimens are evaluated between QC events, and for curve C2, 57 patient specimens are evaluated between QC events. Curve C3 represents a  $1_{2.58s}$  QC rule with 1 QC sample measured at each QC event and 50 patient specimens evaluated between QC events. The false-rejection probability of the QC rules is 0.01 in all cases.

back to 0 for increasingly larger out-of-control error conditions. The maximum value attained by  $E(N_U)$  corresponds to the worst-case situation for the associated QC strategy.

Consider setting performance goals for a QC strategy by specifying a maximum allowable value for  $E(N_U)$ ; for example,  $E(N_U) < 1.0$ . Then the laboratory's task is to design QC strategies (frequency of QC testing, number of QCs measured, QC rules applied) that meet the performance claim. Among alternative QC strategies that meet the performance claim, the laboratory can select the one deemed optimal. For instance, a laboratory might consider choosing the QC strategy that meets its desired quality claim but at the lowest QC utilization rate.

Define QC utilization rate as the ratio of the number of QC samples tested at each QC event divided by the number of patient specimens tested between QC events. There are 2 ways that utilization rate can be modified: change the number of QC samples measured at each QC event (change the numerator) or change the number of patient specimens measured between QC events (change the denominator). Fig. 3 compares the 2 approaches to changing QC utilization while striving to maintain a quality control performance claim;  $E(N_U) < 1.0$ . Curve C1 is the continuous-mode bracketed QC case shown in Fig. 2, which tests 2 QC samples at each QC event and 50 patient specimens between QC

events. Keeping the number of patient specimens between QC events at 50 but reducing the number of QC samples measured at each QC event from 2 to 1 causes the peak  $E(N_U)$  value to greatly exceed the desired maximum value of 1.0 (curve C3). Alternatively, increasing the number of patient specimens between QC events from 50 to 57 while continuing to test 2 QC samples at each QC event results in a peak  $E(N_U)$  value that approaches but does not exceed the desired maximum value of 1.0 (curve C2). Thus, in this case, by increasing the number of patient specimens tested between QC events from 50 to 57, the laboratory could reduce the QC utilization rate by 14% while maintaining its performance claim.

Table 1 demonstrates the impact on QC utilization rate when changing various quality control design parameters for a continuous-mode bracketed QC strategy. For all cases shown in Table 1, the QC strategy has been designed so that the maximum value of  $E(N_U)$  never exceeds 1.0. In each case, the number of patient specimens tested between QC events ( $N_B$ ) has been set to the maximum value that meets the  $E(N_U)$  requirement. The first 2 rows of the table compare the effect of QC rule choice. For the  $\bar{Z}_{1.98}/S_{2.81}$  rule with 2 QC sam-

**Table 1.** The QC utilization rate ( $N_Q/N_B$ ) for various QC strategies that meet the restriction  $E(N_U) \leq 1$ .<sup>a</sup>

| ( $CV_1, CV_2$ ) | QC rule                   | $P_{fr}$ <sup>b</sup> | $N_Q$ | $N_B$ | $N_Q/N_B$ |
|------------------|---------------------------|-----------------------|-------|-------|-----------|
| (3.0, 2.25)      | $\bar{Z}_{1.98}/S_{2.81}$ | 0.01                  | 2     | 57    | 0.035     |
| (3.0, 2.25)      | $1_{2.81s}$               | 0.01                  | 2     | 34    | 0.059     |
| (3.0, 2.25)      | $\bar{Z}_{1.70}/S_{1.96}$ | 0.01                  | 4     | 70    | 0.057     |
| (3.0, 2.25)      | $\bar{Z}_{2.46}/S_{3.48}$ | 0.001                 | 2     | 14    | 0.143     |
| (2.6, 1.9)       | $\bar{Z}_{1.98}/S_{2.81}$ | 0.01                  | 2     | 394   | 0.005     |
| (2.6, 1.9)       | $1_{2.58s}$               | 0.01                  | 1     | 51    | 0.020     |
| (2.6, 1.9)       | $\bar{Z}_{2.46}/S_{3.48}$ | 0.001                 | 2     | 109   | 0.018     |

<sup>a</sup>  $N_Q$ , number of QC samples tested per QC event;  $N_B$ , number of patient specimens tested between QC events;  $E(N_U)$ , expected increase in the number of unacceptable patient results reported during the existence of an undetected out-of-control error condition.

<sup>b</sup>  $P_{fr}$ , false-rejection probability.

ples per QC event ( $N_Q = 2$ ), 57 patient specimens can be tested between QC events while meeting the quality claim that  $E(N_U)$  never exceeds 1.0. For the  $1_{2.81s}$  QC rule (reject if any QC result is more than 2.81 SDs from target), only 34 patient specimens can be tested between QC events and still meet the quality claim.

Comparing rows 1 and 3 of Table 1 shows the impact of changing the number of QC samples evaluated at each QC event. The  $\bar{Z}_{2.46}/S_{3.48}$  rule with 4 QC samples has the same false-rejection probability as the  $\bar{Z}_{1.98}/S_{2.81}$  rule with 2 QC samples (0.01). For the  $\bar{Z}_{2.46}/S_{3.48}$  rule with 4 QCs, however, only 70 patient specimens can be tested between QC events and still maintain the quality claim that  $E(N_U)$  never exceeds 1.0. Thus, the  $\bar{Z}_{1.98}/S_{2.81}$  rule with 2 QCs has a lower QC utilization rate (2/57 vs 4/70) while providing the same level of quality. Comparing rows 1 and 4 shows the effect of decreasing the false-rejection probability of a QC rule. The last 3 rows evaluate the impact of a decrease in the stable analytical imprecision of the measurement procedure by approximately 15%.

## Discussion

A quality control performance measure can serve many roles. It can be used to evaluate and compare alternative QC strategies. It can be used to assist in designing a QC strategy that meets performance requirements specified by the laboratory or others. It can be used by the laboratory to communicate to its customers the level of quality it is able to ensure. Last, it can be used to gain insight into what makes a good QC strategy and why.

The traditional quality control performance measure used by the laboratory to design and evaluate quality control strategies is based on the probability of a QC rule rejection. The probability of a QC rule rejection depends on the choice of QC rule, the number of QC

results used in the rule, and the true state of the process. However, one of the major quality control issues facing today's laboratory has nothing to do with how many QC samples should be tested or what QC rules should be used; rather, it has to do with how frequently QC testing should be performed. The dilemma is that changes in the frequency of QC testing have no effect on the probability of a QC rule rejection. Thus, the performance measure used by most laboratories provides no help or guidance with respect to determining an appropriate QC testing frequency.

One of the goals of this work was therefore to identify and evaluate a quality control performance measure that is affected by changes in QC testing frequency. Additionally, I desired a performance measure that assessed QC performance in terms directly related to the quality of reported patient results. A performance measure that meets both of these goals is the expected increase in the number of unacceptable patient results reported during the existence of an undetected out-of-control error condition,  $E(N_U)$ . This performance measure is sensitive to laboratory testing mode (batch vs continuous) and the nature of the out-of-control error condition (intermittent vs persistent).

For the laboratory testing modes examined here, as the magnitude of the out-of-control error condition increases,  $E(N_U)$  increases, reaches a maximum, and then decreases (Fig. 2). For very small out-of-control error conditions, there is a low likelihood of detecting the error state quickly, but the probability of producing unacceptable results is also extremely low. For very large out-of-control error conditions, the probability of producing unacceptable patient results is high, but the likelihood that the error state is detected the first time it is encountered is also extremely high. Somewhere in between these 2 extremes is the worst case, in which the out-of-control error is large enough that the probability of producing unacceptable results is more

substantial, but the error is still small enough that it is difficult to detect the error state the first time it is encountered. The magnitude of the out-of-control error condition where  $E(N_U)$  peaks and the value of  $E(N_U)$  at its peak will depend on many factors, including laboratory testing mode, total allowable error specification, the nature of the out-of-control error condition, and the power of the QC rules.

The expected increase in the number of unacceptable patient results reported for a batch-mode process with a persistent out-of-control error condition (Fig. 2, curve B) is always greater than for the same batch-mode process experiencing an intermittent out-of-control error (Fig. 2, curve A). The largest difference between the curves is for relatively small out-of-control error conditions where the probability of detecting the error state in the first batch is relatively low. For an intermittent error condition, failure to detect the error state in the first (and only) batch with error incurs no further increase in unacceptable patient results in subsequent batches, whereas for a persistent error condition, repeated failure to detect the error condition increases the number of unacceptable patient results reported.

$E(N_U)$  for a batch-mode process with a persistent out-of-control error condition (Fig. 2, curve B) is also always greater than  $E(N_U)$  for a continuous-mode process with the same persistent out-of-control error condition (Fig. 2, curve C). For a continuous-mode process, because the error condition can occur at any point in the testing stream, the expected number of patient specimens affected by the error condition in the first bracket that contains the error condition will be less than the total number of patient specimens contained in the bracket. For batch-mode processes, all patient specimens in an out-of-control batch are affected by the error condition.

Note that because  $E(N_U)$  has a well-defined maximum, the laboratory can design QC strategies that limit the worst-case performance of the strategy without having to worry about what magnitude of out-of-control error condition might occur. There is no need to compute a critical out-of-control error condition. Note also that if it is reasonable to assume that a laboratory's testing process may be susceptible to persistent out-of-control error conditions as well as intermittent out-of-control error conditions, then the laboratory should design its QC strategies

assuming persistent error conditions, because they are associated with a larger expected increase in number of unacceptable reported results.

Using  $E(N_U)$ , one approach to selecting an optimal QC strategy is to seek the strategy with lowest QC utilization rate among all those that can make a claim that  $E(N_U)$  never exceeds some specified level. As seen in Fig. 3, the ability to fine-tune the QC utilization rate to meet a specified maximum allowable value for  $E(N_U)$  is more readily achieved by changing the number of patient specimens tested between QC events, rather than changing the number of QC samples that are measured at each QC event. Table 1 illustrates some of the many types of QC design decisions that can be addressed in an objective, quantifiable manner in terms of  $E(N_U)$  and QC utilization rate.

In summary, when QC performance is based on the expected increase in the number of unacceptable reported patient results, the impact of the frequency of QC testing can be objectively assessed. Additionally, performance is assessed in terms of the quality of reported patient results rather than the quality of laboratory batches. Finally, QC strategies can be designed to protect against worst-case scenarios that don't require specifying the magnitude of a critical out-of-control error condition.

---

**Author Contributions:** *The author confirmed he has contributed to the intellectual content of this paper and has met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

**Authors' Disclosures of Potential Conflicts of Interest:** *Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:*

**Employment or Leadership:** None declared.

**Consultant or Advisory Role:** Bio-Rad Laboratories.

**Stock Ownership:** None declared.

**Honoraria:** None declared.

**Research Funding:** None declared.

**Expert Testimony:** None declared.

**Role of Sponsor:** The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

## References

1. Clinical and Laboratory Standards Institute (CLSI). Statistical quality control for quantitative measurement procedures: principles and definitions; approved guideline—third edition. Wayne (PA): Clinical and Laboratory Standards Institute; 2006. CLSI document C24-A3; ISBN 1-56238-613-1.
2. Westgard JO. Internal quality control: planning and implementation strategies. *Ann Clin Biochem* 2003;40:593–611.
3. Parvin CA. Estimating the performance characteristics of quality-control procedures when error persists until detection. *Clin Chem* 1991;37:1720–4.
4. Parvin CA. New insight into the comparative power of quality-control rules that use control observations within a single analytical run. *Clin Chem* 1993;39:440–7.
5. Parvin CA. Quality-control (QC) performance measures and the QC planning process. *Clin Chem* 1997;43:602–7.
6. Ye JJ, Ingels SC, Parvin CA. Performance evaluation and planning for patient-based quality control procedures. *Am J Clin Pathol* 2000;113:240–8.