The impact of QC frequency on patient results[1,2]

By John Yundt-Pacheco and Curtis A. Parvin, PhD

Automation. Would it not be great to have the laboratory's quality-control tasks completely automated? No more puzzling about whether to re-run a control specimen or how to evaluate the results. The test system would simply work until it stopped, and it would stop only when something was wrong. What would the set-up parameters look like on such a system? How would you specify the quality goal that the system needed to meet? After giving thought to such questions, it is likely you would realize that conventional QC design parameters, like probability of error detection, are insufficient. A completely automated QC system needs to deal with patient risk, which leads to the following question: What is important in controlling patient risk?

The importance of QC frequency

Frequency of QC plays a major role in managing patient risk. This can be illustrated by considering two laboratories that are using the same QC procedure but with different patient loads:

- Lab A tests two QC specimens every morning and evaluates the results using a two-standard deviation QC rule; Lab B does the same.
- Lab A tests 100 patients per day; Lab B tests 200 per day.
- A grave, undetected malfunction occurs midday that compromises subsequent specimens.
- While Lab A and Lab B are using the same QC procedure, the patient outcomes are not the same: Lab A produced 50 compromised results; Lab B, 100.
- Why is there this difference in the number of compromised results?
- A similar scenario could be constructed with both labs evaluating the same number of patients but with Lab B only testing QC every other day.

The answer is found in the frequency of QC and the relationship between QC specimens and patients. Conventional QC rules originated from batch-oriented testing. There is a direct relationship between the quality of the QC specimen and the quality of the patient specimens in a batch test. The advent of discrete testing has changed this relationship. With a much higher degree of independence between tests, discrete analyzers have a weaker connection between the QC specimens and patients. While the quality of a QC specimen may be acceptable, a subsequent malfunction could compromise patient specimens and remain undetected until it is identified with future QC evaluations. The number of patients that are tested between QC specimens directly affects the risk of producing a compromised result.

Patient risk as a performance metric

Considering patient risk as a performance metric in QC strategy design was first considered by Dr. Parvin and Dr. Gronowski in *Effect of analytical run length on QC performance and the QC planning process*.[3] The average number of patients with an unacceptable analytical error because of an undetected out-of-control-error condition was proposed as a metric for evaluating the efficacy of control strategies. When this metric is applied to error conditions large enough to compromise all subsequent specimens, one-half the patients between QC evaluations are compromised (this assumes that a malfunction is equally likely to occur anywhere — which will put it in the middle on average) for immediate release results.

The average number of patients with an unacceptable analytical error can be reduced by increasing the frequency of QC specimens. The more frequently QC specimens are evaluated, the lower the average number of patients at risk for unacceptably large analytical errors due to an undetected malfunction. Logically, this implies that the way to minimize patient risk is to evaluate control specimens with each patient specimen. While evaluating controls with each patient is typically not practical, the insight that QC frequency plays a major role in patient risk is valuable. When considering how to mitigate large errors, it is a better strategy to intersperse QC specimens between patients rather than group them together.

Bracketed QC strategies

Evaluating a QC specimen with each patient will minimize patient risk, but is not practical — so, what can be done? One possibility is to consider bracketing patient specimens with QC specimens and not reporting patient results until a QC specimen is successfully evaluated. Bracketed QC offers several attractive design features and restores the dependency relationship between the quality of QC specimens and the quality of patient specimens. If evaluation of the closing QC specimen in a bracket determines that no grave, persistent error state exits, then the patient specimens were also evaluated with no grave, persistent error state. Unlike the immediate release case, bracketed QC strategies can be designed that will do substantially better than an average risk of one-half the number of patients between QC specimens.

The expected number of patient results reported with an unacceptable amount of error due to an undetected error condition — $E(N_u)$ — can be used as a design goal. A total allowable error (TE_a) specification is used to determine if the error present in a result is unacceptable. For bracketed QC, the $E(N_u)$ can be computed as the product of the increased probability of a result having an unacceptable amount of error due to an error state (P_E) and the average number of results reported during an error state ($ANP_{reported}$):

$$E(N_u) = \Delta P_E \times ANP_{reported}$$

The probability of a result having an unacceptable amount of error can be computed for any potential error condition. Plotting ΔP_E against systematic error results in a curve that starts at zero when there is no error and eventually increases until it is near one

when the amount of systematic error present is large in comparison to the total allowable error. When high amounts of systematic error are present, almost every result will have an unacceptable amount of error.

The average number of patients reported during an error state has the opposite behavior. When error levels are low, they are very hard to detect — so a large number of patients will be reported before determining that a small error state exists. If a large error state is present, it will be detected by the closing QC specimen; so, no patients will be reported. Plotting the $ANP_{reported}$ against systematic error results in a curve that starts high when there is little or no error and drops as the error increases. When high amounts of systematic error are present, $ANP_{reported}$ drops to zero as the large amount of error will be detected by the closing QC specimen.

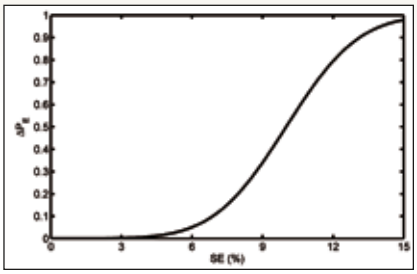

Figure 1. The probability of an unacceptable amount of error (ΔP_E) curve; total allowable error specification (TE_a) is ± 10%.

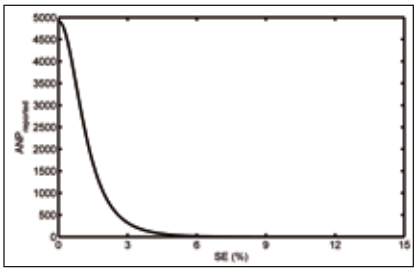

Figure 2. The average number of patients reported ($ANP_{reported}$) curve; a mean/range rule with a probability of false rejection (P_{fr}) of 0.01 is tested every 50 patients.

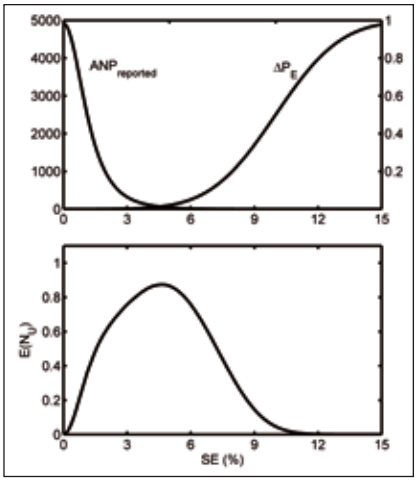

Figure 3. The expected number of unacceptable results — $E(N_U) = \Delta P_E \times ANP_{reported}$; for this case, $E(N_U)$ reaches its maximum at 4.7% systematic error.

The product of the ΔP_E curve and the $ANP_{reported}$ curve is the $E(N_u)$ curve — the expected number of patient results with an unacceptable amount of error as a function of error size.

The $E(N_u)$ curve has a distinct feature — it has a maximum. The $E(N_u)$ curve rises as the amount of error increases until it hits a maximum and then descends until it reaches zero. Intuitively, this is because when there are small amounts of error, the expected number of unacceptable results associated with it will be small — there is little error to cause problems. When there are large amounts of error present, the error will be detected by the closing QC specimen evaluation and results will not be reported. If they were, they would likely be unacceptable. It is somewhere in between where the most damage occurs; the error is large enough to cause problems but not large enough to be immediately detected.

The expected number of unacceptable patient results as a quality goal

From a design standpoint, the $E(N_u)$ is attractive because a QC strategy can be designed for which the $E(N_u)$ will never exceed a maximum that can be specified by the laboratory. For instance, a quality goal can be stated where, in the event of an undetected malfunction, the expected number of unacceptable reported patient results will be less than one. QC strategies can be designed by altering the number of patient specimens between QC evaluations (N_B) and the composition of the QC evaluations (the number of QC specimens, N_Q, and the quality-control rule used to evaluate the specimens) so that the maximum $E(N_u)$ is below the design goal.

For any given $E(N_u)$ goal, there are many (infinite) possible QC strategies that meet the goal. The ratio of the number of QC specimens tested at a QC evaluation to the number of patients tested between QC evaluations (QC utilization rate) and the false rejection rate can be used to select the lowest cost strategy that meets a given goal. The false rejection rate dictates how often the QC rule will indicate a problem when no problem exists — it controls what the $ANP_{reported}$ number is when there is no error in the system. A probability of false rejection of 0.01 is a reasonable place to start. For a given $E(N_u)$, TE_a, and a false rejection rate, we can determine the number of QCs and the number of patients that will achieve the lowest QC utilization rate.

QC strategies that meet a goal of maximum $E(N_U) < 1$					
Analytical Imprecision	Evaluation	False Rejection	Num. QC	Num. Patients	QC Utilization
(σ1, σ2)	QC rule	Pfr	N_Q	N_B	N_Q/N_B
(.13,.18)	Mean/Range	0.01	2	57	.035
(.11,.15)	Mean/Range	0.01	2	394	.005

Table 1. Improving analytical imprecision has a dramatic effect on the amount of QC needed.

In general, QC utilization depends on the TE_a/analytical CV relationship. Improving analytical imprecision has a dramatic effect on the amount of QC needed. In the given example, moving from control SDs of (0.13, 0.18) for levels 1 and 2 to (0.11,0.15) allows the expansion of a bracket from 57 patients to 394 patients.

Summary

The expected number of unacceptable patient results due to an undetected, malfunction — $E(N_u)$ — can be set as a patient-based quality goal. Using the number of patients tested between QC specimens as a design parameter allows one to design QC strategies that meet specified patient-based quality goals. The QC utilization rate can be minimized in a QC design for a given $E(N_u)$. The QC-utilization rate achievable depends on how close analytical imprecision is to the total allowable error.

Currently, **John C. Yundt-Pacheco** is Scientific Fellow at Bio-Rad Laboratories' Quality Service Division in Plano, TX. **Curtis A. Parvin, PhD**, is director of Informatics and Statistics in the Division of Laboratory and Genomic Medicine, Department of Pathology & Immunology, Washington University School of Medicine, St. Louis MO.

References

1. Yundt-Pacheco J. *The Impact of Quality Control Frequency on Patient Results* Presented at the 3rd IFCC Conference on Quality; 2008, Mexico City, Mexico.
2. Parvin CA. *The Impact of Quality Control Frequency on Patient Results* Presented at the AACC Annual Meeting, July 2008, Washington, DC.
3. Parvin CA, Gronowski AN. Effect of analytical run length on quality-control (QC) performance and the QC planning process, *Clin Chem.* 1997;43:11:2149-2154.