Establishing Commonsense-Based Statistical Quality Control Practices

Martín Yago, PhD, and Herminio López-Escribano, PhD

From the Service of Laboratory Medicine, Hospital General de Requena, Valencia, Spain.

**Key Words:** Quality control; Risk management; Analytical quality; Quality control planning

Sustained technological advances have substantially improved the reliability of clinical analytical systems. However, ensuring the analytical quality of the reported results remains a major concern, and statistical quality control (SQC) using stable control materials remains a crucial part of detecting errors caused by excessive variation of the analytical system and thus of preventing reported patient results from containing medically important errors. Traditionally, SQC procedures have been chosen based on their efficiency in detecting critical systematic errors in the measurement procedure (ie, those of sufficient magnitude to cause a significant percentage of test results to contain errors exceeding the quality requirement for the test).

Technological advances have coincided with the introduction of risk management concepts and a reorientation of the clinical laboratory from an activity-based service to a service based on patient outcomes. This has fundamentally changed SQC planning, shifting the focus from analytical systems’ stability to patient safety and from process-centered to patient-centered activity. From this perspective, SQC procedures are no longer selected because of their probability of detecting critical systematic errors in the measurement procedure but for their ability to limit the probability of patient harm due to an out-of-control condition relative to a residual level considered acceptable from the viewpoint of patient safety.

The risk of patient harm due to an out-of-control condition is directly proportional to the number of erroneous patient results reported as a consequence of this condition. The more erroneous results reported, the more likely that inappropriate clinical decisions or actions will be made that cause harm to the patient. Therefore, the probability of patient harm can be reduced via the selection of appropriate SQC procedures that limit the maximum expected number of results that can be reported due to an undetected out-of-control error condition. This amount, usually termed Max E(N UF), can be used to guide the process of SQC planning. Using simple nomograms allows easy estimation of Max E(N UF) for each control procedure from the characteristic σ value of the measurement procedure. Max E(N UF) can also be estimated roughly from the probability of detecting critical systematic errors, as there is a close relationship between this probability and the number of erroneous results reported due to an out-of-control condition.

In this issue of the journal, Westgard and Westgard propose a traditional approach for planning SQC procedures complemented by the use of a patient risk model to determine the frequency of SQC events or the run size needed for the bracketed operation of continuous production processes. The number of patient specimens to be measured between QC events (the run size) is selected so that the Max E(N UF) value is equal to 1. For the bracketed operation of a continuous production process, the authors recommend a “critical control point” mode that provides a high level of error detection at the beginning of operation (called the startup SQC design) and a simpler “monitor” mode that considers the desired reporting interval for the subsequent SQC events. Several tools are provided to facilitate the selection of control rules, number of controls, and frequency of QC events. This selection is based exclusively on the σ value of the measurement procedure.

The authors describe their strategy as “evidence based” in the sense that it is supported by a more-or-less
plausible statistical model but do not provide empirical evidence of their proposal for SQC planning being superior in practice to patient safety achieved in current schemes. Either way, the proposed procedure can provide guidance on the maximum number of patient samples that can be analyzed between QC evaluations so as to report no more than a certain number of erroneous results if an out-of-control condition occurs.

However, a critical step in SQC planning is to establish precisely how many erroneous test results will be reported due to an eventual out-of-control condition without posing an unacceptable risk to patients. To establish this number, it is necessary to examine the analytical system to identify and evaluate potential causes of failure and to determine the risk of harm they represent for patients.

The risk of harming patients depends not only on the number of erroneous results reported in each out-of-control event but also on the frequency with which these situations occur (ie, the reliability of the analytical system), the likelihood that the erroneous result is used to make erroneous decisions that cause harm to patients, and the severity of the resulting harm. Given that these factors generally vary between tests, each test will have its own acceptable Max E(N_{UF}) value that depends on the influence of the test on decisions regarding patient care and the reliability of the analytical system that generates the result.

The Max E(N_{UF}) value can vary by changing the number of patient samples analyzed between QC events (as suggested by Westgard and Westgard in the aforementioned study) and also by using different control rules and numbers of controls per QC event. Note that the value selected as an objective for Max E(N_{UF}) has important economic consequences because its value is inversely proportional to the number of patient samples that can be analyzed between QC events. Setting Max E(N_{UF}) as 1 instead of 2, for example, involves at least doubling the cost of control procedures because the number of QC events must also be doubled. A low value for Max E(N_{UF}) may not be achievable if the capability (σ) of the measurement procedure is low and requires the analysis of a disproportionate number of controls; such a value might be unnecessary if the analytical system is highly reliable.

The maximum acceptable number of erroneous results reported due to an out-of-control condition in a highly reliable analytical system is greater than that for a less reliable system because in the former, this condition occurs infrequently. Therefore, improvements to the analytical process that lead to a reduction in the rate of failure (eg, frequency of calibration, lot-to-lot changes control, or better maintenance procedures) will allow the use of SQC procedures that are more lax and less expensive. SQC planning is a dynamic process that can take advantage of improvements in the reliability of the analytical system independently of the σ of the measurement procedure.

Failure modes should also be taken into account when planning SQC. For instance, the probabilities of failure during startup and during monitoring of the analytical process usually vary between measurement procedures. Some are unstable, and their analytical performance worsens within a few hours, while others are stable for months but have a risk of failure at critical points such as changing reagents. QC schedules should be designed in accordance with these behaviors (eg, SQC procedures with greater ability to detect errors should be used at the stages in the production process with greater probability of failure).

Another factor that determines the number of erroneous results that a laboratory can afford to report during an out-of-control condition is the risk that each erroneous result represents for the patient. The reporting of an erroneous result can create a hazardous situation for patients and can lead to harm if the result is used to make a wrong clinical decision. The patient risk will depend on the likelihood that an inappropriate decision will be made as a result of an erroneous result and the severity of any ensuing harm. An estimate of this risk can be made for each test by considering its clinical use and the health care context in which it is used. The higher the potential severity of harm arising from an erroneous patient result, the lower the acceptable probability that the harm will occur and the lower the number of erroneous results that can be reported during an out-of-control condition without putting the patient at excessive risk.

In the era of patient-centered medicine, evidence-based SQC strategies should by definition take into account the impact of an analytical error on patient safety, considering the testing process in all its complexity. Statistical models provide valuable information for SQC planning but may convey a false sense of certainty when they do not include a complete view of the process. An SQC system based solely on the σ value of the measurement procedure can lead to the adoption of control strategies that do not respond to the patients’ safety needs and therefore should be discouraged.

SQC planning is more a decision-making process based on knowledge of the analytical system and common sense than a mathematical exercise based on statistical models. It involves making fine-grained decisions to ensure that the selected SQC strategies are simple,
cost-effective, and fit for purpose. The estimation of several parameters to be considered in SQC planning (eg, the frequency of failure, the probability that an erroneous result leads to an inappropriate decision, and the severity of any resulting harm to patients) includes a strongly subjective component, although it is possible to treat the parameters quantitatively. However, these parameters can dramatically affect the adequacy and cost of the selected control procedures, and the subjectivity of their estimation should not be a reason to ignore them when planning SQC. As stated by the British philosopher Carveth Read, it is sometimes “better to be vaguely right than exactly wrong.”

References