Limitations of Quality Control in Physicians' Offices and Other Decentralized Testing Situations: The Challenge to Develop New Methods of Test Validation

Daniel M. Baer¹ and Richard E. Belsey²

Traditional quality-control methods are effective for detecting systematic error caused by deterioration of reagents or instruments, but ineffective for detecting sporadic error, which is more likely to occur in low-volume testing environments. Decentralized testing performed by individuals without formal laboratory training has a high potential for sporadic errors. New methods for validating test results, used each time a test result is produced, should replace current quality-control procedures. Under the rules of CLIA '88, manufacturers and the US Food and Drug Administration have an opportunity to develop new approaches to test validation.

Quality-control (QC) testing has been used since the 1950s as a means of ensuring that the variability of laboratory test results is within allowable limits. QC, as currently practiced in clinical laboratories, had its origin in manufacturing, when industry became concerned with the consequences of producing parts that were unusable because they no longer met production specifications. High-volume manufacturing machines performing repetitive functions tend to deteriorate over time (e.g., stamping dies tend to wear or chip), resulting in product variation that makes the product no longer suitable for its intended use. The cost of discarding or reworking such parts was so high that a system was developed to test selected pieces from a production run to be sure that the parts being produced were within acceptable limits. If a piece that was QC tested was not within allowable limits, another piece was tested to be sure that there was a problem. If the problem was confirmed, the production system was shut down until the problem was located and corrected.

In the laboratory, we know that analytical system components tend to wear out or change with time: e.g., light bulbs and optical surfaces in photometers can develop coatings that shift spectral wavelengths, reagents can deteriorate, and mechanical components such as pumps can wear, changing the volumes of reagents or samples measured. Because the true laboratory product, the patient’s care result, could not easily be used as a quality index, the laboratory adaptation of industry’s QC techniques involved analysis of a pseudospecimen containing known amounts of constituents (a control sample). Whenever a batch of patients' samples is processed, a control sample is also tested to determine whether the QC result is within acceptable limits so that the analytical system is, presumably, functioning normally. When the QC result is outside allowable limits, the operator can take corrective action to avoid the reporting of a possibly erroneous patient's care result. Conditions that lend themselves to detection by this kind of QC are as follows:

• changes that are progressive, starting small and becoming larger until they are no longer acceptable;
• changes that are predictable;
• changes that continue over a group of successive production runs;
• processes for which all aspects can be tested by the QC technique; and
• testing systems that have high-volume throughput.

The high-volume testing environment allows the ratio of QC testing to patients' test output to be low while providing reasonable assurance that a problem will be identified before an erroneous result is produced. The use of a sampling technique assumes that testing conditions at the time of QC sampling are similar to the conditions present before and after the selected QC sample inspection was produced.

Current laboratory QC techniques focus on the analytical system's production run and not on the operator of the system. The system's automation, together with the operator's training, have allowed the operator to be considered a neutral factor in QC testing. The operator's role in the QC process is to

• interpret the QC result to decide whether the patients' care results can be released or whether a possible problem exists,
• retest a control specimen to see whether the QC result is still outside acceptable limits, and
• undertake problem identification and correction, if appropriate.

Traditional QC is, consequently, effective for discovering systematic errors. The rules of Westgard et al. (1), for example, are designed to discover shifts, biases, and excessive variability but not sporadic errors that occur only in patients' care samples. An erroneous patient's sample result due to a transient clot in a pipettor tip might, for example, be caught by the operator because of a marked change from previous results, if they existed, but would be “invisible” to the usual QC techniques.

¹ Laboratory Service, Veterans Affairs Medical Center, 3710 SW U.S. Veterans Hospital Rd., Portland, OR 97207.
² Department of Pathology, Oregon Health Sciences University, Portland, OR 97201.

Received May 11, 1992; accepted September 15, 1992.
The error would not be apparent if such previous results were not available.

Over the years, the quality of laboratory testing has had major improvements, in part because of the use of these QC techniques. Now, important changes in the practice of laboratory medicine prompt us to reexamine the role and adequacy of current QC techniques in validating and assuring the quality of test results used in decisions affecting patients’ care. These changes include improvements in analytical test systems, which have made them less expensive, more stable, and simpler to operate than in the past, thereby supporting the transfer of complex technology to a work force with limited technical training.

At the same time, new provisions of CLIA '88 (Section 493.1201) permit manufacturers, with Food and Drug Administration (FDA) approval, to specify acceptable QC measures for their medical testing devices. This provision provides an opportunity for manufacturers and the FDA to develop innovative approaches that have the potential to improve the quality of laboratory testing, particularly in the hands of individuals with limited technical training, e.g., in such sites as physicians’ office laboratories, hospital nursing units, nursing homes, and limited service laboratories. Here, we examine the limitations of traditional QC techniques in decentralized settings, and suggest new, more effective approaches to the validation of test results.

Limitations of Traditional Quality Control

Testing systems developed for use by individuals with limited technical background (e.g., in physicians’ office laboratories, hospital nursing units, and nursing homes) generally operate in a low-volume testing environment. Reagents are generally very stable and are prepared as single, totally contained testing units that are frequently supplied individually wrapped. The following examples illustrate some of the kinds of errors that can occur with such testing systems:

• A glucose meter used on a hospital nursing unit is used in succession for three patients. Following the proper operating procedures, the nurse cleans the optical window before performing the first test. The second patient’s blood contaminates the window during the test, but a valid test result is produced. The optical window of the meter is partially covered with blood at the time of the third patient’s test, and an erroneous result is produced.

• A nurse on a nursing unit prepares to perform a capillary blood glucose test on a patient. He removes a test strip from the vial and places it on the bedside stand; before he is able to perform the test, he is called away to care for a critically ill patient. Meanwhile, a nursing aide comes to give the patient a bed-bath. In the course of the bath she puts a wet washcloth on top of the reagent strip, introducing moisture and causing deterioration. Later the nurse returns to the bedside and performs the test, using the deteriorated strip.

• A home health nurse makes a visit to a patient for whom determination of prothrombin time is needed. In his bag he has three reagent cassettes for the bedside capillary blood prothrombin time test, which he took from the refrigerator when he left the health center. He uses only one of the cassettes and forgets to return the other two to the refrigerator when he gets back to the office. A week later, he uses one of the cassettes that he had carried around in his bag for a week, having exposed it to ambient temperatures warmer than the manufacturer’s recommended storage conditions.

• A receptionist in a physician’s office is also responsible for performing laboratory tests. The office laboratory uses a test system that uses individual foil-wrapped test slides. She unwraps a test slide and places it in the instrument to perform a test on a venous blood sample. Before she can perform the test, she receives a phone call about a complex billing problem, has to register several patients, and performs other tasks for the physician. An hour and a half later, she returns to complete the testing process.

• Later in the day the same receptionist is performing an immunological test for Group A beta hemolytic streptococci. Part way through the procedure, the telephone rings; because of the interruption, she forgets what step she is performing and omits the addition of one reagent.

The likelihood that such sporadic errors or mistakes will occur in testing done by individuals with limited technical background is, as suggested in these examples, probably higher than the likelihood that such errors would occur in conventional laboratories. We believe that the operator and his or her interfaces with other components of the test system are the areas that are most prone to testing errors. In some hospitals, where bedside capillary blood glucose testing is done by many nurses, the institution has assigned responsibility for QC testing of the instrument and reagents to a centralized QC coordinator. Unfortunately, random operator errors, such as those described above, are rarely detected by these techniques. In the hospital setting, it is, according to the Joint Commission on Accreditation of Healthcare Organizations, the institution’s responsibility to monitor the reliability of individual operator’s results and take corrective action in the case of operators performing below pre-set norms (2). Conventional laboratory QC techniques are unlikely to detect sporadic operator errors when they occur in the testing of patients’ care samples. Some of the problems we think are most likely to occur in these settings include the following:

• timing errors;

• reagent or system exposure to stressful environmental conditions such as excessive temperature or humidity;

• errors due to reagent contamination or dirty optics;

• the use of reagents that are outdated or have deteriorated prematurely;

• miscalibration of reagent sets;

• errors through omission of a reagent or use of reagents in the wrong order; and

• errors of specimen or reagent sampling and dispensing.

Table 1 illustrates systematic and sporadic errors that affect the production of reliable test results.
New Approaches Necessary

We believe that the traditional QC approach is ineffective for detecting the sporadic errors that have the highest likelihood of occurring in the setting of near-patient, low-volume testing. QC techniques need to be developed, adapted, or modified to address potential problems (including those listed above) to validate the reliability of results used in decisions affecting the care of patients. To us, this means that the QC check needs to be an integral part of the analysis of each patient's test sample; in conventional terms, the batch size for near-patient testing needs to be 1, and the QC technique needs to include monitoring for operator error as well as for analytical system error.

Some manufacturers have incorporated "procedural" or "process" checks into test systems, especially those that use membrane-based immunological techniques (e.g., pregnancy tests and group A beta hemolytic streptococci tests). These checks provide the operator with an indication that the procedure was performed correctly or, conversely, that the result may not be reliable and should not be reported.

Other manufacturers have introduced electronic or optical system checks in their systems. These take the form of colored cards or chips used in capillary blood glucose meters or cassettes that simulate a colored result in some chemical and coagulation systems. Although these devices are useful for problem-solving once it has been shown that a problem exists, we believe that they are being confused with true QC testing, which they are not: only the latter provides the operator with an indication that a test result may be unreliable. Over the years, the variety of descriptors of quality management of laboratory testing—quality assurance, quality control, quality management, and proficiency testing, among others—has confused lay managers and laboratory professionals. We, therefore, propose more descriptive terminology to help simplify communication about the techniques used to ensure the reliability of patients' care results on a day-to-day basis. We recommend the following terminology.

Validity check—a procedural check built into the reagent unit of the test system: it detects exposure to damaging environmental conditions and is carried through all steps of the procedure to ensure that the proper reagents are used in the correct order and that the proper timing sequences are followed.

Quality-control test—traditional QC with a QC sample (pseudospecimen) that is similar in matrix and constituents to the material being tested and that responds to reagents and testing systems in the same way as does a patient's sample.

System check—a mechanical, electronic, or optical device that tests all or part of the optical, mechanical, or electronic component of the testing system but does not involve a QC sample (pseudospecimen) or a patient's sample.

The Nature and Importance of Validity Checks

To be useful to the intended user (the operator performing the test), validity checks will need to

- be test specific,
- be used to document the normal functioning of the system at significant clinical decision values for each test. If necessary, a second validity check should be included so that the normal operation of the system is documented at both values,
- screen for substances that would commonly produce erroneous results (a validity check for a potassium test, for example, should include a screen of the sample for hemoglobin as an indicator of cell hemolysis, a significant confounding factor in this determination),
- be a unique indicator in noninstrumented systems (such as the colored dot or plus or minus signs found in some of the membrane-based immunological systems),
- include clear instructions that the result cannot be released if the validity checks are not favorable.

Validity checks will be most useful at the time that the patient's care result is produced. Absence of the validity check can then serve as an immediate indicator, warning the operator that the result may not be valid and should not be released. In instrumented systems, the absence of favorable validity checks should prevent the instrument from reporting the test result. A favorable validity check could be indicated by simply displaying or reporting the result.

Validity checks can effectively deal with the operator as a significant source of analytical error. They can also address some specimen issues affecting specific tests. They cannot yet address other important aspects of the testing process that can affect the reliability and interpretation of the test result: patient preparation, specimen collection and identification, and sample protection and preparation.
Our recommendations were developed in the context of low-volume testing by individuals with limited or no technical training. The language of CLIA '88 now provides an opportunity to develop new approaches for validating the reliability of results produced in these settings on a day-to-day basis. In addition, the personnel requirements under CLIA '88 are expected to result in a shift towards employment of less technically trained individuals in moderately complex and highly complex laboratories as well. The concept of integral validity checks is just as appropriate in these settings. Such an approach has the potential for improving the quality of testing of patients' samples and lowering the cost of QC, which depends on the use of QC materials, uses additional reagents, and requires substantial effort from technologists.

The Role of the FDA and Manufacturers

Manufacturers and the FDA now have an opportunity to develop innovative methods for validating the reliability of test results. The Premarket Approval (PMA) and 510(k) procedures specified by the FDA should require a complete discussion by the manufacturer of measures for detecting sporadic errors by users who have incomplete technical training. This discussion should include the opportunities for error in the system, including the exposure to adverse environmental conditions, timing, contamination, and dirt, and the opportunities for errors in measurement of sample.

Manufacturers and the FDA must assume, for testing performed in patient-care areas by individuals with limited technical training, that a substantial fraction of users will not use the test system as directed. Manufacturers need to indicate to the FDA what fail-safe mechanisms and validity checks are built into the system to prevent or detect conceivable adverse events or misuse of the system.

As part of the PMA and 510(k) approval process, the manufacturer should be obliged to present clinical trial data verifying that the validity checks and other safeguards have been effective in preventing the reporting of erroneous results when test analysis was done by the intended users (e.g., nurses doing bedside testing). Sources of error—including the patient's preparation, specimen collection and identification, and sample handling and preparation—that can cause an erroneous or uninterpretable test result to be produced should be discussed in simple, nontechnical language. This information should fit on a single card that can be prominently displayed on an instrument or on an essential part of a noninstrumented system.

References