# Quality Laboratory Performance

**Ronald H. Laessig, Ph.D.[1]***
**Sharon S. Ehrmeyer, Ph.D.[2]**
**State Laboratory of Hygiene[1] and Medical Technology Program[2]**
**Department of Pathology and Laboratory Medicine[1,2]**
**University of Wisconsin-Madison**
**Madison, Wisconsin**

***Presenting Author**

The narrator of “Star Trek” defines space as the last frontier. In laboratory practice research, achieving "quality laboratory performance," in the 1995 meaning of the term, may also be the last frontier. Defining quality laboratory performance, however, assumes that one can define quality. The now-famous video, "Reviewing the Quality Moment," explores the growth and development of the quality concept through the eyes, words and works of giants such as Deming, Juran, Crosby, Youdin and others. The video concludes with each person defining quality. To the average viewer, the dissimilarity and apparent discontinuity in the definitions is almost comic relief to the heavy subject discussed for two hours. It does, however, illustrate an important truth - - quality is an elusive concept. Quality does not have any single, universally accepted definition. Nowhere is this more evident than when we are challenged to define, measure and ultimately improve the "quality of laboratory performance."

Also in 1995, clinical laboratory testing, once the sole purview of centralized laboratories, is now dispersed to sites in physicians' office labs, the ICU and CCU, ER, near patient, clinics, hospitals, commercial megalaboratories and patients' homes. Quality assumes both a universal definition (customer defined need) and practical laboratory-specific definition (complying with a myriad of regulatory requirements).

For our purposes, Crosby's definition of quality is the best: "Quality has much in common with sex. Everybody is for it... everybody feels they understand it...everyone thinks execution is only a matter of following natural inclinations...most people feel that all the problems in this area are caused by other people..."[1] Perhaps this is best illustrated by examining an evolutionary definition of clinical laboratory quality:

Figure 1 chronicles the past indicators of quality, or at least who defined quality, associated with evolving quantitative laboratory tests; they were relatively simple. They also track the history of our lab profession. Starting from Shewhart's work[2], Levey and Jennings[3] suggested that quality limits should be set at mean $\pm$ 2 standard deviations (SD). Westgard later called this the $1_{2s}$ rule.[4] This led to multiple other rules, $1_{3S}$, $2_{2S}$, etc. and later multiple rule applications. Application of quality rules, especially those based on the laboratory's own, self-established mean value and SD, as later required under CLIA '67[5], is really not an accuracy-based approach but is a means of monitoring the stability of the analytical

**An Evolutionary Definition of Quality in Laboratory Testing**

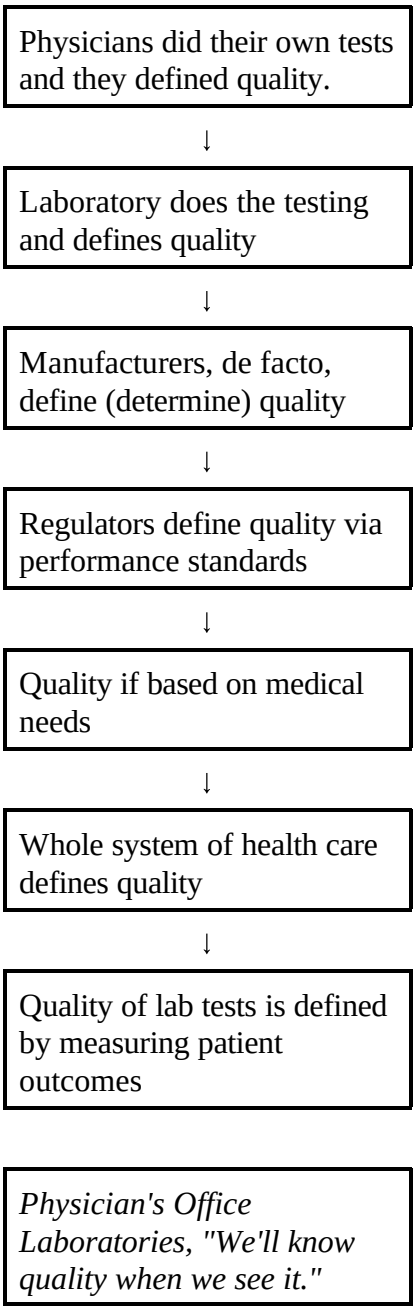


Figure 1

system. The $1_{2s}$ rule, coupled with the admonition, "if the control is out, rerun it," enabled laboratories to adequately monitor the performance of 1950's methods typically run on a Coleman Junior Spectrophotometer™ or later AutoAnalyzer I™. By today's standards, these methods and techniques with poorly defined calibrations and large SDs would be considered inferior, poor quality, and probably unsuited for routine use. The point is that the processes to define and monitor quality and other factors have allowed us to improve.

Curiously, this mean and SD approach to quality, which essentially evolved on an empirical basis, is not inconsistent with that of Dr. Deming.[6] Deming stresses that plotting data is an essential first step to an intelligent, systematic approach to achieving or (continuously) improving the quality associated with a process. That process could be a laboratory test method. The plotting itself is primarily to help the user determine the stability of the system. While clinical laboratories prefer the terms "in control" or "out of control" to "stable" and "unstable," once the stability of the process is established, then according to Deming, quality improvement efforts can occur. Interestingly, the requirement of CLIA '67 that laboratories recalculate the mean and SD on a monthly basis and use it to monitor and evaluate quality in the following month was actually a form of continuous quality improvement (CQI). If the method's performance improves, for whatever reason, the improvements are incorporated into subsequent monitors (quality indicators, rules, etc.) of the process. The requirement is implicitly carried forward in CLIA '88,[7] although laboratory directors seem to be given some latitude in applying the quality mandate. Under CLIA '67, however, clinical laboratories departed from Deming's theory because a formal commitment to CQI as part of managing the testing process was absent. This was included in CLIA '88. Efforts of the College of American Pathologists (CAP) that focused on external Quality Assurance (i.e., proficiency testing) could also be viewed as a commitment to CQI. As participants “got better," the group SD became smaller; thus, the standard of quality was to be in the central 95%. The scheme was not planned to be CQI--but it was. In the era pre-CLIA '88, the true quality question almost never asked was "Is the observed SD good enough?" The focus was on the measurement process itself -- not improving it. Perhaps, inherently, we believed that the level of quality was so inferior that universal improvement was needed.

The first traditional aspect of quality targeted for improvement was precision or, more correctly, imprecision. Laboratories of the 50's and 60's had to accept what the method could give. The pressure to reduce variation, based on the assumption that less variation or imprecision in the measurement process is inherently better, came first from the manufacturing sector, where claims of smaller SDs were viewed, and marketed, as indicators of quality, and later from interlaboratory (i.e., megapool) comparison programs. CAP's QAS program and manufacturers' programs, such as those originally offered by Hyland, Dade and General Diagnostics, allowed laboratories to intercompare, not only means, but also SDs on the same lot numbers of quality control products. Peer comparison then, as now, was a powerful incentive to everyone to improve. It worked then; it works now!

The second traditional aspect of quality is

## QUALITY TIME LINE - A VERY SELECTIVE MEMORY

**1912**
Folin-Wu glucose test

**1912**
Testing begins to migrate to "big" labs

**1945**
Belk-Sunderman launch proficiency testing
CAP begins formal PT

**1950**
Levy and Jennings - QC charts
First laboratory inspections
Automation comes to clinical laboratories
Radin - "What is a standard?"

**1966**
Medicare/Medicaid

**1967**
CLIA'67 passed
NCCLS launches concept of voluntary standards

**1976**
FDA establishes labeling regulations
CDC hosts Congress on reference methods
FDA product class standards come and go
Scandals in drug, Pap testing push Congress to legislate for labs

**1988**
CLIA amendments passed
NCCLS holds CLIA Congress
JCAHO incorporates TQM into inspections

**1992**
CLIA's first final CLIA regulations

**1995**
HCFA formally turns CLIA toward TQM

Figure 2

accuracy. Accuracy, in many respects, is more difficult to monitor and improve. Milestones in the process were initiating the use of standards in clinical labs. Dr. Radin discussed this concept in 1965, and again in 1995, with an announced plan to review the topic again in the year 2025. [8,9] Accuracy assessment progressed through two tracks. One was the acceptance of controls with assigned values (i.e., "Versatrols"™). These and similar products which we could not accept as true standards were dubbed "calibrators" by the “Gang of Four” (Boutwell, Mother, Vanderlinde and Laessig while working on a product class standard for the Food and Drug Administration (FDA) in the 70's). The other, in 1947, was the introduction of interlaboratory proficiency testing (PT) by Belk and Sunderman.[10] Later, when true standard values were replaced with reference serum-based, empirically assigned, target values in calibrator products, PT (that is, comparisons to peers or "peer group means") became the standard means for assessing accuracy. As anti-intellectual as the "democratic" approach to defining the right answer (i.e., accuracy) seems, on its face, one thing can be said: "it works." The laboratory community has been working on reproving Beck and Sunderman’s 1947 findings for nearly 50 years.

In the last part of this presentation, I will cite some actual current data from Wisconsin's HCFA-approved PT program to demonstrate that this empirically derived process works in even relatively unsophisticated laboratories.[11] Dr. Sunderman would be justified in saying, "been there, done that!"

We can also view Quality Laboratory Performance as a laboratory quality timeline (Figure 2). It not only includes the points discussed above, but also includes two other sentinel events which bring forth another questionable principle, namely, that of assuring quality performance through legislation. The enactment of the Medicare/Medicaid legislation in the early 1960s [12] was followed closely by the CLIA '67 regulations which applied to a limited, but significant, subset of laboratories. Basically, the theory behind the legislation was that if proven quality practices, such as 1) running controls, 2) participating in PT, 3) having a procedure manual, etc., can be identified, simply mandating them will ensure that the practices are implemented in regulated laboratories, and the attendant quality follows. Needless to say, the concept has been debated.

While CLIA '67 focused on minimal performance standards and compliance with mandated quality control/quality assurance practices as a means of achieving the desired end result (i.e., promoting "high quality" laboratory performance); CLIA '88's regulations take a different tack. These regulations incorporate a "total quality management" approach which has two aspects: (1) quality is defined by focusing on the customer's needs or expectations (customers being patients and clinicians), and (2) the laboratory is asked to design and implement a quality assurance (QA) system suitable to achieving a level of performance commensurate with its definition of quality. This is described as mandating the "what" or "outcome" and not the "how" it is to be achieved. The Joint Commission on Accreditation of Health Care Organizations (JCAHO) originally endorsed this "TQM" approach in their inspection standards before the CLIA '88 regulations were promulgated.

Numerous approaches have been attempted to defining quality in the clinical

laboratory, initially in terms of imprecision and later in terms of total error. Both approaches sought to develop a quality standard by asking clinicians how much error, in terms of inaccuracy and imprecision, could be tolerated. Though customer focused, these attempts were unsuccessful. Of late, clinicians have been asked to react, under a given set of patient circumstances, to hypothetical changes in patient test values of varying magnitude. This approach to quantifying total allowable error as the size of change which precipitates a reaction shows some promise. The advantage of this is that it allows clinicians to think in terms of the patient and the way they practice using our laboratory results. When pressed, quality experts stress defining quality in terms of the customers' needs. Customers can, in this case, be physicians or patients.

JCAHO, HCFA (CLIA) and CAP inspection processes currently focus on the laboratory director's responsibility to improve quality. Specifically, the laboratory director, under CLIA's section §493.1218, is required to define, design and implement a QC system adequate for the laboratory. To quote but one example of the CLIA language: "...the laboratory shall determine the number, type and frequency of controls to ensure results adequate to patient needs." Implicit in the CLIA'88 regulations is the idea that the laboratory director must establish a quality control (QC) program, not just follow a collection of mandated QC practices, to ensure that the laboratory produces medically useful data. CLIA '88 allows the laboratory director to decide, in the absence of specific mandates or rules, what, quantitatively, constitutes adequate performance. It empowers the inspector to look for evidence of this process, but clearly the director is responsible for devising the strategy and defining the quality parameters. Surprisingly, CLIA's position does not differ greatly from that long endorsed by CAP. CAP's inspection process always has placed a great deal of credence on the judgment of the laboratory director.

CLIA '88, unlike its predecessor CLIA '67, also broke new ground by defining minimum standards of interlaboratory performance (accuracy and precision) essentially expressed as total error based on the proficiency testing (PT) criteria for specific regulated analytes. As some of our previous work has demonstrated, these PT performance standards can be translated into minimum, intralaboratory performance requirements.[13,14] For example, acceptable PT performance for glucose is defined as target value $\pm$ 10%. In the absence of any significant bias (inaccuracy), this translates into an intralaboratory precision requirement of approximately 4% or less. A laboratory with a 4% CV and no bias will rarely fail glucose PT. The presence of bias reduces the portion of the error budget assignable to imprecision. Our colleague, Jim Westgard, cautions that this approach assumes the presence of a stable (in control) operating system.[15] If the laboratory procedure is out of control or beginning to fail, the 4% error budget is too large to ensure successful PT performance and, by implication, quality patient testing. It is generally accepted that in writing the PT regulations, HCFA did not project an error budget for the laboratories, but rather used an empirical approach based on past inter-laboratory performance, primarily in CAP PT surveys.

CLIA '67 by relying on a simple $1_{2s}$ rule for QC, as suggested by Levey and Jennings, inherently assumed the acceptance and perpetuation of the current status as the standard of performance, i.e., quality. It

failed to incorporate the concept of "customer need" into the quality specifications. CLIA '88 gives us a mandate to incorporate the concept of medical usefulness into our laboratory's approach to designing a quality assurance system. Once an error budget can be defined as a quality specification, approaches like Westgard's QC Validator™ software allows for establishing performance standards based on the desired outcome. In this case, the outcome is a QC approach which ensures a designated level of accuracy and precision. This is, in our opinion, the best that we as a profession have to offer today. As laboratorians look to the future, however, the word "outcome" takes on a different meaning.

**Research Into Indicators of Laboratory Performance:**

To date, anecdotal research, that is, reporting on what is observed, has demonstrated a continuous improvement in quality of laboratory results. Quality in this case is defined as improved accuracy, decreased imprecision and laboratory test methods that generally demonstrate higher sensitivity and specificity. A major quantum improvement in the quality of laboratory results also has been in the area of test availability, manifesting itself as extensive menus of tests, even for small, less sophisticated laboratories, as well as greatly improved turnaround times (i.e., near patient testing). These are descriptors of test performance, they carry the implicit assumption of quality, or at least process improvement, but actually do little to speak to the fundamental definition of quality, which must ultimately focus on "fitness for intended use."

At this conference in October 1995, under the general heading of *Research Indicators of Quality*, one is obligated to report on "the experiment" in progress. In 1994, all laboratories engaged in moderate or high complexity testing in the U.S. were mandated to participate in PT for a very significant number of regulated analytes. Beginning in 1995, PT performance was used as a means of evaluating and documenting intralaboratory performance as a condition of continuing to hold a valid "CLIA" certification. HCFA steadfastly has indicated that the purpose of PT is not punitive nor is it designed to shut down laboratories. Instead, its purpose is focused on education and laboratory improvement. "The experiment" can be illustrated by citing some data from the Wisconsin HCFA approved PT Program. The 1994 data, in Table 1a, represent largely physician office laboratories (POLs) engaged in PT for the first time.[11] The 15% failure rates for cholesterol in shipment #1 is followed by a similarly high failure rate, 12%, for all laboratories in shipment #2. Referring to the lower portion of the chart (Table 1b), however, the number of laboratories failing both events is only 3%. The implication is that of the 15% that failed initially, four out of five were able to improve their performance and not fail in the second event. One certainly is free to speculate as to what happened, but in these POLs, the most likely scenario is that once alerted to a failure, the director, typically not a pathologist nor clinical chemist, etc., took some appropriate action. That appropriate action could be as simple as calling the instrument manufacturer or a laboratory colleague (pathologist, chemist or medical technologist) for assistance. These data again demonstrate the premise inherent in Belk and Sunderman's report from 1947, "When conscientious laboratorians are informed of a situation

**TABLE 1a**
**1994 Performance Data**
**WSLH's Chemistry Proficiency Testing Program**
"2 of 5" Analyte Failures (%)

| Analyte | 94-1 | N | 94-2 | N | 94.3 | N |
|---|---|---|---|---|---|---|
| Albumin | 9 | 223 | 7 | 223 | 3 | 232 |
| ALT (SGPT) | 6 | 565 | 5 | 563 | 8 | 552 |
| Alkaline Phosphatase | 4 | 442 | 4 | 433 | 3 | 433 |
| Amylase | 5 | 415 | 4 | 414 | 4 | 389 |
| AST (SGOT) | 6 | 621 | 4 | 603 | 7 | 599 |
| Bilirubin, total | 8 | 482 | 7 | 460 | 5 | 465 |
| Blood gas $PCO_2$ | 3 | 173 | 1 | 172 | 1 | 176 |
| Blood gas pH | 2 | 173 | 1 | 172 | 2 | 176 |
| Blood gas $PO_2$ | 2 | 173 | 2 | 172 | 3 | 176 |
| Calcium | 10 | 314 | 8 | 314 | 5 | 323 |
| Chloride | 8 | 281 | 8 | 274 | 8 | 285 |
| Cholesterol, total | 15 | 1094 | 12 | 1083 | 11 | 1057 |
| Cholesterol, HCL | 10 | 728 | 10 | 704 | 8 | 663 |
| CK isoenzymes | 1 | 80 | 0 | 79 | 0 | 81 |
| Creatine kinase | 5 | 335 | 2 | 316 | 2 | 306 |
| Creatinine | 8 | 825 | 8 | 804 | 7 | 795 |
| Glucose | 12 | 1099 | 7 | 1080 | 8 | 1060 |
| Iron, total | 1 | 84 | 5 | 89 | 4 | 95 |
| LDH | 6 | 325 | 3 | 309 | 6 | 319 |
| LDH isoenzymes | 6 | 17 | 0 | 17 | 0 | 17 |
| Magnesium | 8 | 172 | 6 | 172 | 1 | 175 |
| Potassium | 7 | 915 | 7 | 898 | 6 | 882 |
| Sodium | 17 | 479 | 18 | 476 | 16 | 481 |
| Total protein | 8 | 297 | 7 | 291 | 7 | 292 |
| Triglycerides | 8 | 952 | 5 | 935 | 4 | 907 |
| Urea nitrogen (BUN) | 14 | 851 | 14 | 836 | 12 | 815 |
| Uric acid | 6 | 786 | 4 | 768 | 5 | 738 |

**TABLE 1b**
**1994 Performance Data**
**WSLH's Chemistry Proficiency Testing Program**

| Analyte | 94-1 to 94-2 (2 consecutive) | 94-2 to 94-3 (2 consecutive) | 94-1 to 94-3 (2 out of 3) | 94-1 to 94-3 (2 consecutive or 2 of 3) |
|---|---|---|---|---|
| Albumin | 2.7 | 0.0 | 0.9 | 3.4 |
| ALT (SGPT) | 1.6 | 0.4 | 1.3 | 2.5 |
| Alkaline Phosphatase | 0.7 | 0.5 | 0.2 | 1.4 |
| Amylase | 0.5 | 0.3 | 0.0 | 0.8 |
| AST (SGOT) | 1.2 | 0.5 | 1.3 | 2.7 |
| Bilirubin, total | 1.1 | 1.1 | 0.2 | 1.9 |
| Blood gas $PCO_2$ | 0.0 | 0.0 | 0.0 | 0.9 |
| Blood gas pH | 0.0 | 0.0 | 0.0 | 0.0 |
| Blood gas $PO_2$ | 0.0 | 0.6 | 0.0 | 0.6 |
| Calcium | 1.9 | 1.5 | 1.9 | 4.0 |
| Chloride | 1.8 | 1.1 | 1.4 | 3.5 |
| Cholesterol, total | 3.3 | 2.2 | 2.6 | 6.3 |
| Cholesterol, HDL | 2.1 | 1.5 | 0.9 | 3.8 |
| CK isoenzymes | 0.0 | 0.0 | 0.0 | 0.0 |
| Creatine kinase | 0.0 | 0.3 | 0.0 | 0.3 |
| Creatinine | 1.5 | 1.6 | 0.9 | 3.3 |
| Glucose | 2.3 | 1.4 | 1.6 | 4.1 |
| Iron, total | 1.1 | 0.0 | 0.0 | 1.1 |
| LDH | 0.6 | 0.3 | 0.9 | 1.3 |
| LDH isoenzymes | 0.0 | 0.0 | 0.0 | 0.0 |
| Magnesium | 1.7 | 0.0 | 0.0 | 1.7 |
| Potassium | 1.2 | 1.2 | 1.0 | 2.2 |
| Sodium | 4.8 | 4.8 | 4.6 | 10.0 |
| Total protein | 2.4 | 0.7 | 1.0 | 2.7 |
| Triglycerides | 1.2 | 0.6 | 1.2 | 2.1 |
| Urea nitrogen (BUN) | 4.2 | 2.8 | 2.2 | 6.9 |
| Uric acid | 0.9 | 0.7 | 0.7 | 1.8 |

needing improvement, and failing an event in PT is such a situation, appropriate, effective action almost always follows -- with or without regulations." Another view might be that health care professionals, laboratorians and/or primary care providers want to, and given the opportunity, will do the right thing.

**Quality Laboratory Performance -- Alternative Benchmarks:**

A justifiable criticism of laboratorians, and thus far, perhaps the authors of this paper, could be that of taking the traditional view of quality (Figure 3). Quality improvement and measuring quality have heretofore been focused on the testing process itself. The inherent assumption, unproven and untested, is that "improvements in test quality descriptors lead to improvements in health care." In a true quality driven health care system, one which will focus on demonstrating positive outcomes, these traditional approaches fall short of the mark. Considering the current status of health care, two general, outcome-based, measurement models can be projected: Model #1 has the ill or injured patient encountering the health care system and being restored to health. As one colleague (somewhat cynically) described it: "If the patient walks away from the hospital, the outcome is positive and the process must have worked." In Model #2, the patient benefits from prevention and early detection and does not need an in-depth encounter with the system. Both models assume that "some process" has worked correctly to promote, create or ensure a positive outcome and, implicitly, that laboratory testing is an important part of that process.

Considering Model #2, using Wisconsin's Newborn Screening Program (NSP) as an example, the implicit value of testing can readily be demonstrated. As shown in Table 2, the NSP screens approximately 70,000 newborns a year at a cost of $2.6 million for all laboratory work, initial followup and dietary or therapeutic regimens for phenylketonuria (PKU) and hypothyroidism. Screening 70,000 children in Wisconsin yields approximately 22 positive cases of either PKU or hypothyroidism (HYT). Projecting a 40- year lifetime in an institution for these 22 infants, undetected cases will cost $50,000 per year per child, suggesting that the direct costs to the State of Wisconsin will be $44 million over their lifetime. Simplistically, newborn screening results in a net cost savings of about $41.4 million per year. On the basis of cost alone, this clearly is a favorable outcome. In these cases, the quality of life outcome for the 22 individuals growing up as essentially normal children, and for their families, is also a positive outcome. The quality of the testing process is clearly a contributor to a very favorable, positive outcome. Conversely, a failure (a false negative, a missed child) is a disaster. In a more quantitative definition of quality, the quality of the newborn testing process could be assessed by defining a target of 100% sensitivity and 100% specificity. The degree to which this is achieved could also be an outcome-based measurement of quality. As we will discuss momentarily, outcome is more difficult to assess -- to measure success, the process would have to include tracking down the baby, initiating treatment and demonstrating an improved quality of life. Tough to do, but not impossible.

Under Model #1 (the sick person encounters system - returns to health), finding a means for using outcome measurements to assess the quality of routine testing is orders of magnitude more difficult.

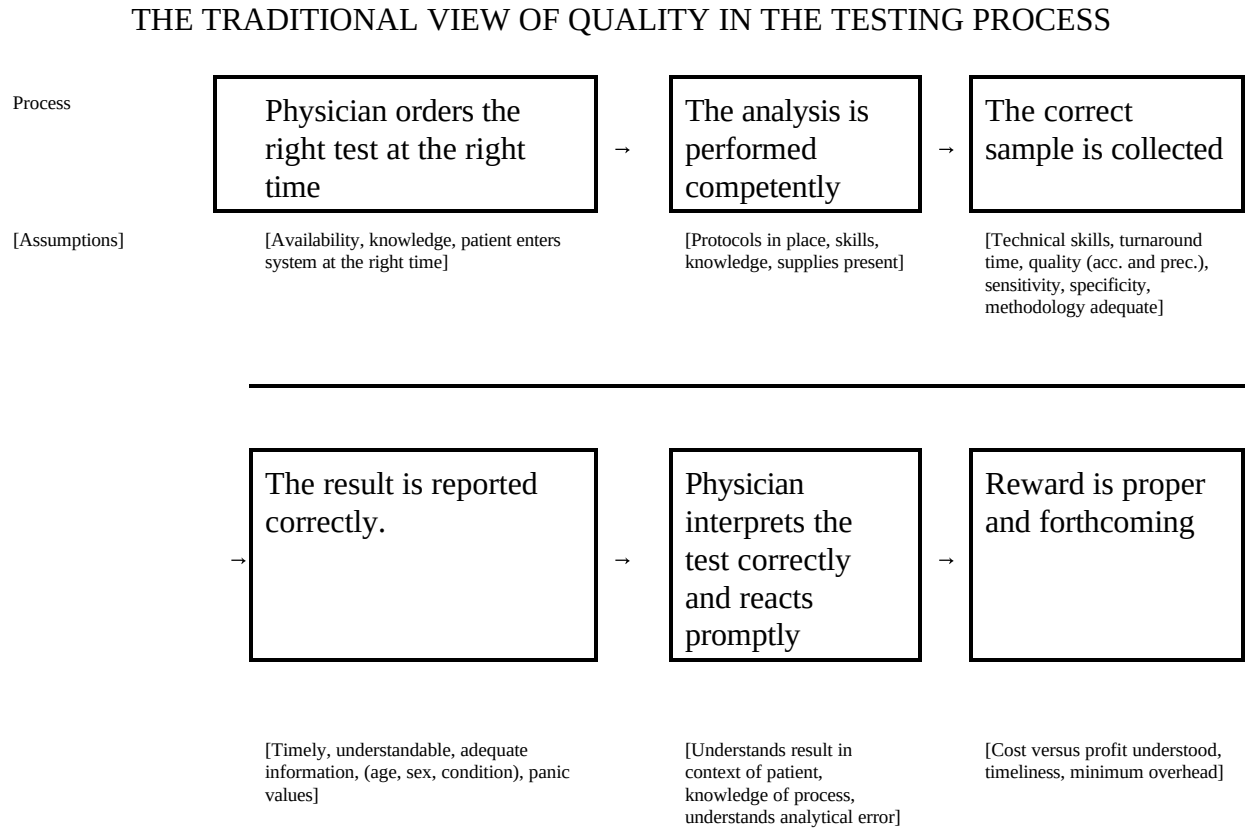


Figure 3

Unless a way is found to do precisely this, however, laboratory testing in the health care system of the future will be facing a continuous barrage of assaults from those in competition for the health care dollar. In addition, the laboratory will need to defend and demonstrate the value of its contribution as new policies are developed and implemented. The model of the ill or injured patient interfacing with the system and being restored to health can be quantified in terms of favorable outcomes for the patients. However, laboratorians need to do much better-- we must demonstrate and quantify that laboratory's testing played a proactive role in achieving the favorable outcome. Even tougher to do.

In Model #1, these assessment criteria for demonstrating laboratory test quality come immediately to mind: (1) essential to making the diagnosis, (2) altered the course of treatment, (3) demonstrated need for treatment, or (4) demonstrated that intervention or further treatment was not necessary. These data presently are not available, except in very anecdotal form. As large numbers of individual cases are studied, retrospectively by an objective observer, the role and effect of laboratory test processes should be demonstrable. In addition, aggregating large numbers of cases will enable such outcome-based studies to define

WISCONSIN NEWBORN SCREENING PROGRAM

| Test | Incidence | Tests/year in Wisc | Cases Detected |
|---|---|---|---|
| PKU | 1/20,000 | 70,000 | = 3.5 cases |
| HYT | 1/3,800 | 70,000 | = 18.42 cases |
| | | | 21.92 = 22 cases/year |

| | | | |
|---|---|---|---|
| 70,000 × 21.50 | = | $1,505,000 | Laboratory component |
| 70,000 × 16.00 | = | $1,112,000 | Formula/therapeutic agent component |
| | | $2,617,000 | Cost to system per year |

SIMPLE COST - BENEFIT COMPUTATION

22 cases × 40 institution years × $50,000 year/case = $44,000,000

NET SAVINGS TO SYSTEM

$41,383,000/year

Table 2

optimized treatment algorithms and protocols. These can and should include laboratory testing. This can and must be done by evaluating the entire process; the laboratory is not an independent entity in the role. This process will: (1) take a truly dispassionate observer and evaluator, and (2) require studying very large numbers of cases. The unfortunate situation in today's capitated health care systems is that treatment protocols, including laboratory testing, are focused and evaluated against cost containment objectives. General apprehension is prevalent among the laboratory community and health care professionals that cost containment protocols are somewhat arbitrary and rarely backed by sound, detailed, scientific data. There are exceptions, of course.

What we have described above is only Phase 1 of a study that defines quality of laboratory performance in the context of outcome-based measurements. Assuming that Phase 1 could be successfully undertaken, Phase 2 is even more difficult, but even more necessary. Once treatment norms and protocols have been described, further studies will be necessary to determine issues, such as optimum timing of tests, the effects of various levels of traditional quality (tied up in the concept of total allowable error), and use of reflexive testing, on outcomes. In short, Phase 2 will focus not only on "what" and "how often" tests are to be run, but more precisely on test use and interpretation. Using actual case studies, questions such as the appropriate time to order tests, reflexive testing, protocols based on optimum combinations of test results and patient-related information could be

developed. The goal is finding the optimum overall approach, which also means ascertaining the most optimum effective use of the laboratory's capabilities. The process must focus on criteria which are outcome measurement-based not only on the restored health of the patient, but also the optimum path to achieve the favorable outcome. This could mean fewer tests; it could mean more tests!

Admittedly, such studies are difficult, complex and extremely time-consuming. If, on the other hand, laboratorians cannot demonstrate that the testing processes improve patient outcomes, it will be difficult to justify our continued existence in a future health care system. The newborn screening models do point the way and demonstrate what can be expected. However, taking the next step, that is, looking at outcomes which include the more routine tests, is a bigger challenge. Our late Wisconsin colleague, Dr. Charles Altschuler in Milwaukee in the 1970's, invented the Programmed Automated Laboratory Information (PALI) system to demonstrate the potential value of computer-ordered laboratory tests. The system proved that it is possible to make diagnoses of insidious disease, optimize care, and reduce the length of stay in a hospital. This approach was clearly outcome based. Unfortunately, the users of the data, in one case surgeons, were not interested in making a new diagnosis and in another case, the administrators viewed shortening hospital stays as a negative consequence of the PALI testing process. How times have changed!

In summary, the determinants of the quality of laboratory testing have moved from undefined, to the province of physicians, to laboratorians, to manufacturers, to regulators, and now, whether we like it or not, to the whole health care system. In a very real way, they have moved back to physicians. The attributes of quality, when laboratory performance was, by today's standards, generally poor, were smaller SDs, smaller analytical biases, high sensitivity and high specificity. Today, the attributes of quality must be defined in terms of the contribution of the laboratory's processes to a favorable patient outcome within the context of the entire health care process that includes laboratories as one integral component. As laboratorians, we are no longer in a silo looking up at our own small portion of the sky. We are now only one player in a much bigger system of players where the quality of our contribution must be, and will be, viewed in the context of contributions to a truly favorable outcome.

**References**

1. Crosby PB. *Quality is Free*. New York: American Home Library, 1979.

2. Shewhart, WA. *Economic Control of Quality by the Manufactured Product.* Van Nostrand, New York, NY, 1931.

3. Levey S, Jennings ER. The use of control charts in the clinical laboratory. *Am J Clin Pathol.* 1950; 20:1059-1066.

4. Westgard JO, Groth T, Aronsson T, et al. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. *Clin Chem*. 1977; 23:1857-1867.

5. Department of Health, Education and

Welfare, Public Health Service, Centers for Disease Control. Clinical Laboratory Improvement Act of 1967. Part F, Title III, PHS §353.

6. Deming WE. *The New Economics for Industry, Government, Education.* Cambridge, MA: MIT, Center for Advanced Engineering Study, 1993.

7. U.S. Department of Health and Human Services. Medicare, Medicaid and CLIA programs. Regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA). Final Rule. *Federal Register*. 1992;57:7002-186.

8. Radin N. What is a standard? *Clin Chem.* 1967;13:55.

9. Radin N. How standards were developed. *Clin Chem.* 1995; 41:S270.

10. Belk WP, Sunderman FW. A survey of the accuracy of chemical analysis in clinical laboratories. *Am J Clin Pathol.* 1947;17:853-861.

11. Ehrmeyer SS, Burmeister BJ, Laessig RH. Laboratory performance in a state proficiency testing program: what can a laboratorian take home? *Clin Immunoassay.* 1994; 17:223-230.

12. Federal Health Insurance for the Aged: Regulations for Coverage of Service of Independent Laboratories. Public Health Report 13. Washington, DC: Department of Health, Education, and Welfare; 1968.

13. Ehrmeyer SS, Laessig RH, Leinweber JE, Oryall JJ: The 1990 Medicare/CLIA final rules for proficiency testing: minimum intralaboratory performance characteristics (CV-Bias) needed to pass. *Clin Chem.* 1990;36:1736-40.

14. Laessig RH, Ehrmeyer SS, Leinweber JE. Health Care Financing Administration's new proficiency testing rules: characterizing long term performance as "successful," "probation" and "suspended." *Arch Pathol Lab Med.* 1992;116:770-776.

15. Westgard JO, Burnett RW. Precision requirements for cost-effective operation of analytical process. *Clin Chem.* 1990;36:1712-1716.