Proficiency Testing 1995-2000: Educational Tool, Quality Control Tool, Management Tool, or a Regulatory Tool?

Sharon S. Ehrmeyer, Ph.D.*
Ronald H. Laessig, Ph.D.
Department of Pathology and Laboratory Medicine
University of Wisconsin-Madison
Madison, Wisconsin

***Presenting Author**

Abstract: Proficiency testing (PT) is all of the above; numerous citations demonstrate the efficacy of each application. The real question is: "What is PT to be, circa 2000, in the context of CLIA'88?" PT was conceived by and for laboratory directors as an educational tool based on interlaboratory performance comparisons. As a quality control (QC) tool, PT has not reached its potential. The drawback is the lack of timeliness in reporting. More recently, with mandated director's review of responses to PT failures, it is an integral part of management and quality assurance. In 1995, PT is the lynchpin of the CLIA'88 regulatory process. The data, from laboratories participating in PT for the first time in 1994, demonstrate both its ability to affect performance improvement, as well as, PT's known limitations. The opportunity in the future for PT depends on redefining its mission and role. To transform PT to a new level of effectiveness as a QC tool, timeliness must be addressed. Minimizing turnaround-times (TAT) through technology will allow a quantum improvement in the value of the PT process. If the 60 day TAT, currently mandated under CLIA'88, becomes 60 seconds or less, a whole new quality paradigm is possible. The challenge of implementing this new vision is only to dare to dream!

Introduction

Interlaboratory proficiency testing has nearly a 50 year history in U.S. clinical laboratories.[1,2] During this time, its role as an educational, quality control, management or regulatory tool has been under continuous re-examination. For the many analytes requiring on-going PT participation and evaluation under CLIA'88, the question has been answered unequivocally: it is a regulatory tool.[3] However, even with CLIA'88 the strongest critics concede, and evidence today seems to support, PT still has an opportunity to fulfill educational, quality control and management roles. While it might be debated whether each role is enhanced or diminished by the regulatory focus, no one can deny that PT makes an impact on today's clinical laboratories.

Proficiency testing probably was best described by Forney as the "...distribution of (identical) unknown samples to laboratories for the purpose of determining the ability of laboratory personnel to achieve the correct analysis.[4]" Forney's definition incorporates the evaluation of accuracy through the interlaboratory assessment process. The most often used criterion for evaluation of accuracy is some form of the consensus "right answer." However, regardless of the criterion used to define "good" (acceptable) and "bad" (unacceptable) PT performance,

laboratories, by comparison to peers, have an opportunity to assess the quality (accuracy) of their performance.

When used in a regulatory context, the fundamental premise of PT is that if a laboratory performs acceptably in PT, it also analyzes patient samples correctly. The question that continues to plague the laboratory community, however, is related to the reliability of PT as an indicator of intralaboratory quality. Certainly at a minimum, PT is a means of assessing at least one form of accuracy. It is generally agreed that PT does not measure precision with any degree of reliability. However, the underlying question relates to the suggestion that PT samples are treated differently than routinely processed patient samples. Most recently, with the enactment of CLIA'88 and specified acceptable performance, the question of the validity of the evaluation criteria has further compounded the reliability issue.

PT as an Educational Tool

As originally envisioned by Belk and Sunderman, the PT process was educational, used to apprise laboratory directors as to when their analytical processes varied from that of the collective, wisdom of the group.[5] Belk and Sunderman maintained that competent laboratory directors would take appropriate corrective action when a problem was identified. Curiously under CLIA'88, PT failures mandate that the laboratory director develop a plan of correction. In 50 years, PT has not strayed too far from the original concept.

The College of American Pathologists (CAP) began offering a limited number of voluntary, interlaboratory (i.e., PT) surveys on an organized basis as early as 1947.[2] The obvious success of the PT process in improving laboratory quality led to a proliferation of voluntary interlaboratory testing programs by CAP, professional societies, and state and even municipal health departments during the 40's and 50's.[2,6] With the enactment of CLIA'67, PT became a mandated, but still primarily a self-directed, improvement process for large hospital and reference laboratories.[7,8] The rationale for mandating PT participation was that if a laboratory director could use data from PT as a means of self-assessment of quality, regulators could use the information for the same purpose. While CLIA'67 was relatively vague on what constituted acceptable levels of performance, and even left the PT providers to interpret data in terms of satisfactory or unsatisfactory performance, CLIA'88 does not. The step from a self-assessment to a minimum standard, performance requirement took place when the CLIA'88 regulations, as proposed by CDC and HCFA, included specific performance criteria.

PT as a Quality Control Tool

Intralaboratory error consists of two components - imprecision and inaccuracy. Laboratories assess imprecision through daily intra-laboratory QC activities, leaving the inaccuracy component to be assessed by some other means. For most laboratories accuracy is assessed through PT. PT, especially in Europe, also is called external QC. In the U. S., QC tends to focus more on standard deviations, or imprecision, and concentrates on achieving stable performance. The statistical mean determined in the QC process monitors drift and is not used to assess accuracy. Under CLIA'88, however, PT is linked to the broad area of QC. The clear implication is to make imprecision assessment and accuracy

monitoring a part of the QC process.

An offshoot of PT is related to peer comparison data. Regionalized QC programs such as those originally made popular by Hyland, Dade and General Diagnostics in the 1970's and, now CAP's QAS program, compare performance for laboratories analyzing the same lot of QC material.[9] Similar to PT, these programs, through mean and standard deviation comparisons, t-tests, Youdon plots, etc., offer an accuracy assessment in addition to a daily evaluation of imprecision.

PT as a Management Tool

Some of the ground-breaking differencesincorporated into CLIA'88 focus on PT as a management tool. Belk and Sunderman originally decided that PT should alert the director to potential problems within the laboratory; CLIA'88 requires the director to review PT data, document the review, and also to approve the remediation of any problems identified by the process. As a management tool, PT, originally and today, primarily provides data to determine: 1) the relationship of a given laboratory to peer laboratories, usually peer laboratories using the same methodology; 2) the robustness of methods (good versus poor quality) by assessing the amount of variation and pass rate among the peer group, and 3) the relationship between methods. The latter is the topic of Dr. Laskey; method comparisons are overlaid by the problem of matrix effects which have long plagued PT programs attempting to understand methodology differences.[10]

PT- The Regulatory Lynchpin of CLIA '88

CLIA'88 broke new ground regulating all laboratories, approximately 160,000. A large number of these (70-80,000) perform moderate and high complexity testing requiring PT participation. The drafters of the CLIA regulations should perhaps be applauded for the wisdom of their approach. As managers, they have devised a method for someone else to provide and grade the samples and then to send, in electronic form, the final results to HCFA. HCFA uses the data to accomplish its goal of assessing participant performance. This is not a statement of malevolent intent; it is a statement to acknowledge successful management practice, tempered only by the question as to whether PT should be used for regulation at all.

Critical Assessment of PT Performance Under CLIA'88

All laboratories, including physicians' office laboratories (POL), were required to enroll in a HCFA approved PT program by January 1994. Before 1994, PT participation for POLs was voluntary and results were used for educational purposes. PT data available from Wisconsin's HCFA-approved PT program and California's program indicate that significant performance problems exist, particularly for laboratories participating in PT for the first time.[11-13] In Wisconsin's program for example, 15% of all POL participants failed cholesterol on the initial survey in 1994 and 12% failed on the second survey. However, the good news was that only about 3 % of the laboratories failed both surveys. This indicates to us that most laboratories experiencing problems corrected them by the second survey. The data also show, however, that a large number of laboratories have rather marginal performance. The cholesterol performance limits are relatively generous, i.e., target value plus or minus 10%. Dr. Karen Nickel

reports similar findings in California, where 29% of all CLIA-certified POLs failed, that is, were unsuccessful in two out of three successive PT surveys, for at least one analyte.

Wisconsin's preliminary data from the second year of POL participation indicate continued improvement; further results from the California program are not yet available. Basing our assumption on these data, however, we would project that Belk and Sunderman will again be proved correct and performance will continue to improve. Interestingly, while PT data indicate a definite need for laboratory improvement and demonstrate that PT is an effective mechanism to achieve improvement, the wise men in Washington are seriously thinking of abandoning the process, at this point in time, for POLS. This is clearly brilliant thinking!

PT in the Year 2000 and Beyond

Objectively evaluating the PT process, one can see some positive attributes. Criticisms,however, include the fact that PT is expensive, time-consuming and disruptive to laboratory service. In addition, the number of samples are too small for meaningful interpretation, the process is flawed in that "good" laboratories sometime fail and "bad" laboratories may pass, the evaluation criteria may not be appropriate, and the PT sample matrix affects results. While the list of criticisms goes on, the lack of timeliness between analysis and result evaluation is, perhaps, the biggest drawback, preventing the achievement of PT's full potential as a quality assessment activity.

Making a quantum improvement in timeliness is, in our view, critical to the future of PT. As visionaries, we must not be afraid to dream. The information superhighway is in place. The possibility of reporting and evaluating PT results in "real time" opens the door to a whole new paradigm, one with tremendous opportunity for both laboratories and regulators. The PT process then could combine the attributes of both QC and PT into a single process, enhancing the cost effectiveness of the quality assurance activity. Appropriate computer algorithms, along with new designs of products, open the possibility of assessing multiple aspects of quality, all in real time and on-line, including accuracy, precision, linearity, reportable range, sensitivity, specificity, and method comparisons.

The vision for PT in the future should not be limited to what it can be; we should instead focus on what we want it to be. We must have the courage to achieve the dream.

References

1. Sunderman FW. The origin of proficiency testing for clinical laboratories in the United States. In *Proceedings of Second National Conference on Proficiency Testing.* Information Services, Bethesda, MD. 1975.

2. Dorsey DB. The evolution of proficiency testing in the USA. In *Proceedings of Second National Conference on Proficiency Testing.* Information Services, Bethesda, MD. 1975.

3. 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. Fed Regist. 1992; 57:7002-186.

4. Forney JE, et al. Laboratory evaluation

and certification. In *Quality Assurance Practices for Health Laboratories.* SL Inhom, ed., American Public Health Association, Washington, DC. 1977.

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

6. Skendzel LP, Copeland BE. An international laboratory survey. *Am J Clin Pathol. 1975* 63:1007-1011.

7. Department of Health, Education and Welfare, Public Health Service, Centers for Disease Control. Clinical Laboratory Improvement Act of 1967. Part F, Title 111, PHS §353.

8. 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.

9. Lawson N, Rej R. The Validity of Proficiency Testing, and Performance in Proficiency Testing. *In Proceedings of Frontiers in Laboratory Practice Research,* Centers for Disease Control and Prevention, Atlanta, GA. 1995.

10. Laskey F. Achieving and assessing acceptable analytical performance: The challenge of matrix effects. In *Proceedings of Frontiers in Laboratory Practice Research,* Centers for Disease Control and Prevention, Atlanta, GA. 1995.

11. Burmeister BJ, Lanphear BT, Ehrmeyer SS, et al. Actual 1994 performance data from Wisconsin's HCFA-approved proficiency testing programs: Implications forlaboratories and regulations. *Clin Chem.* 1995;41:S210 [Abstract]

12. 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.

13. Auxter S. POL exemption debate to begin on Capital Hill. *Clin Lab News.* 1995;21:1,12.