Summary of Workshop 1: Proficiency Testing

Facilitator: Paul Bachner, MD
Department of Pathology
University of Kentucky Hospital
Lexington, Kentucky

DLS Liaison: Shahram Shahangian, Ph.D.

Key Questions:

1) Does proficiency testing provide a reliable measure of actual laboratory performance?
2) How can the validity and utility of proficiency testing be enhanced?

Abstract: This workshop reviewed the benefits and drawbacks of current proficiency testing models and considered alternative approaches to proficiency testing. The latter included multi-program characterization of laboratory performance and performance evaluation based on retrospective reference method analyses of specimens previously tested by participating laboratories. Workshop participants also endorsed further study of "hybrid" quality assessment systems in which proficiency testing and quality control activities were blended. The participants concluded that proficiency testing is an important but incomplete measure of laboratory performance, and that multi-programmatic characterization of laboratory performance and restructuring of current proficiency testing models should be actively pursued.

The Workshop Session addressed two key thematic questions: (1) Does proficiency testing (PT) provide a reliable measure of actual laboratory performance? and (2) How can the validity and utility of PT be enhanced?

We began with a review of the seminal publication of Belk and Sunderman,[1] which described a survey performed 50 years ago and inaugurated the modern era of proficiency testing. Testing for common chemical and hematological analytes by 50 volunteer laboratories in Pennsylvania showed sufficient variation in results to place patients into very different clinical management scenarios. Subsequently, laboratory directors who were questioned about possible reasons for poor performance suggested, in order of frequency, poorly-trained and inadequate numbers of technicians, lack of understanding between pathologists and staff, poor institution floor space, and other miscellaneous reasons. Workshop participants observed that much progress has occurred in the intervening years and that the potential for degradation (and improvement) of performance is always present in the process components of laboratory practice.

Four invited presentations (published elsewhere in these Proceedings) developed the substrate for subsequent discussions. The first of these by Dr. Robert Rej of the New York State Department of Health (NYSDH) addressed the question of whether PT measures natural test performance. He stressed that the only components of the complete testing cycle that are

mirrored by PT are analysis and calculation, whereas non-analytic steps are reflected partially or not at all. Furthermore, process inconsistencies and analytic substrate (matrix) differences may contribute to a potential for enhanced or degraded performance. Dr. Rej also reported on a new initiative at the Wadsworth Center (NYSDH) dubbed "Retro-PT" in which hand-carried samples are analyzed by laboratories and subsequently re-analyzed by reference methods at the Wadsworth Center. These studies demonstrated relatively small biases with little potential for clinical impact and that the analyte-specific bias magnitude was similar to those noted in NYSDH PT testing programs.

Dr. Noel Lawson discussed multi-programmatic characterization of laboratory performance based on College of American Pathologists (CAP) data. He presented new data showing that relative PT performance is consistent over time (particularly poor performance) and across PT programs (linearity), performance is better in laboratories participating in interlaboratory PT programs and that similar biases are noted. His data also demonstrated that performance is better as a function of the length of time that a laboratory has been enrolled in PT testing and in laboratories that have been CAP inspected and accredited. These observations were noted in several programs and in multiple studies performed in 1984-86, 1988-90, and 1991-94.

Dr. Fred Lasky of Johnson & Johnson Clinical Diagnostics, Inc., spoke of performance problems associated with the manufacture of PT materials, specifically the impact of "matrix effects" detected in PT surveys but not noted with fresh patient samples. Dr. Lasky presented several recommendations and conclusions including that 1) fresh specimens are better for assessing performance, 2) manufacturing process improvement is possible, and 3) the relative benefit of such changes is questionable in view of the greater need to concentrate on improving non-analytic performance factors.

The final speaker, Dr. Sharon Ehrmeyer, of the University of Wisconsin, summarized the state-of-the-art and identified a "vision" for the future of PT not constrained by present programmatic limitations. Her presentation emphasized that current PT is expensive, time-consuming, disruptive, not timely, and provides incomplete identification of performance problems, even in "good" laboratories.

Based on these presentations, the lively and extensive discussions of participants resulted in several broad areas of agreement that are summarized as follows:

- A NEED TO RESTRUCTURE PT TO BETTER ASSESS NON-ANALYTIC PERFORMANCE: Although recognizing the difficulty of achieving this recommendation, the need to move toward this goal will be accentuated by the increasing utilization of point-of-care testing and the decentralization of testing. The potential for rapid, real-time data transfer through the information highway will encourage emerging regional and local initiatives that try to blend PT and quality control efforts within integrated health care delivery systems.

- WE ARE ON THE THRESHOLD OF HYBRID PT AND QUALITY CONTROL (QC) CONTROL SYSTEMS and that it will be important to validate and compare performance

between different models of what workshop participants provisionally identified as "inter-community PT/QC." To support this trend it will be necessary to look beyond current PT models as well as explore "alternative" QC practice that will be more appropriate to emerging point-of-care technology and instrumentation.

- A CONTINUING ROLE FOR CURRENT NATIONAL PT to provide traceability for "hybrid" PT/QC in local/regional as well as evolving national networks based in large hospital systems and commercial laboratories.

- THE CONTINUING BENEFIT OF CURRENT PT to provide a useful estimate of the state-of-the-art, a system to monitor individual laboratory quality improvement and an early-warning system for problem identification, to satisfy the need for independent assessment, as well as to meet regulatory requirements and considerations of public accountability.

- THE VALUE OF MULTI-PROGRAM CHARACTERIZATION as presented by Dr. Lawson was endorsed and it was stated that hybrid PT/QC could be viewed as a variant of multi-program characterization of laboratory performance that could incorporate blind and split sample testing as components.

Additional important observations made by Institute participants during the presentation of the workshop proceedings to the Plenary session were the serious potential for compromise of the educational and quality improvement role of PT by continuing emphasis on the regulatory and punitive aspect of PT. Such an emphasis will divert PT program providers from designing challenges that will encourage laboratory improvement in order to avoid excessively high failure rates. A related observation was the need for PT program providers to recognize the participating laboratory as their prime customer, rather than government regulatory agencies.

In summary, the Workshop participants concluded, in response to the key questions that were posed, (1) that PT is an important indicator of laboratory performance, but only one of other (QC, quality assurance, inspection, personnel standards, patient test management) partial and incomplete measures of laboratory quality, and (2) that the validity and utility of PT should be enhanced by an increased emphasis on multi-programmatic characterization of laboratory performance and by the restructuring of PT (and hybrid PT/QC programs) to better reflect the
entire testing cycle.

References

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