Received 11.29.06 | Revisions Received 12.1.06 | Accepted 12.5.06

Proficiency Testing Performance in Physician's Office, Clinic, and Small Hospital Laboratories, 1994–2004

Daniel C. Edson, MS, MT(ASCP),[1] Laura D. Massey, MA, MBA, MT(ASCP)[2]

([1]American Proficiency Institute, Traverse City, MI, [2]Missouri State University–West Plains, West Plains, MO)

DOI: 10.1309/6EV4QT8119DU7LXE

Abstract

Background: CLIA 88 was fully implemented for profiency testing (PT) by 1994, but, to date, few studies have examined the long-term impact of PT on performance in laboratories.

Methods: Failure rates for selected chemistry, hematology, and microbiology analytes were monitored periodically from 1994–2004 to evaluate proficiency test performance.

Results: Failure rates for chemistry and hematology analytes declined significantly during the 10-year period. Failure rates for microbiology analytes also declined but remained above 5% in 2004 for positive genital/GC cultures, positive urine cultures, and Gram stains. Performance of chemistry and hematology tests improved significantly during the study period. Microbiology analytes also showed significant improvement, but certain tests remained problematic.

Conclusion: The data indicates that statistics for unsatisfactory laboratory performance may fail to detect significant problems.

Among other changes, the Clinical Laboratory Improvement Amendments of 1988 (CLIA '88) expanded the requirement for participation in a proficiency testing (PT) program to include all laboratories performing moderate- or high-complexity testing on patient samples. As a result, thousands of previously-unregulated laboratories in physicians' offices and clinics were required to enroll in proficiency testing programs by 1994, the year CLIA '88 was fully implemented. Physician's office and clinic laboratories have now been performing proficiency testing for 13 years, but, to date, few studies have examined the long-term impact of proficiency testing on performance in these laboratories.

To evaluate how PT has impacted performance in these laboratories, we examined data from one proficiency test provider, American Proficiency Institute (API). The API began offering proficiency testing to physician's office and clinic laboratories in 1993, and its clients now include more than 13,000 physician's office, clinic, and small (less than 100 beds) hospital laboratories. Since previously-unregulated laboratories comprise most of API's client base, the data from these PT events seem well suited for evaluating performance trends in this segment of the clinical laboratory industry. To this end, we monitored proficiency testing failure rates for selected analytes from 1994 to 2004 to assess whether performance improved in these laboratories. This data, which was presented as poster sessions at various symposia on quality issues in laboratory practice from 1995 to 2005, is summarized and analyzed here.

Materials and Methods

Proficiency test samples were manufactured for API for chemistry (Consolidated Technologies, Inc., Austin, TX), microbiology (Microbiologics, Inc., St. Cloud, MN), and hematology (Streck Laboratories, Omaha, NE). The samples were assembled in kits at API's headquarters in Traverse City, MI, and then distributed to laboratories via 2-day delivery service. Results were analyzed using API's proprietary software, stratified into appropriate peer groups, and then evaluated according to criteria prescribed by the Centers for Medicare and Medicaid Services (CMS).

For this analysis, we defined "failure" as an unacceptable result for an individual sample, as determined by CMS criteria. The failure rates in **Table 1** and **Table 2** are the percentages of unacceptable results for an analyte during a year. For example, out of 39,598 responses for cholesterol in 2004, 1,248 (3.2%) were unacceptable ([1,248 / 39,598] × 100 = 3.2%) (**Table 1**). Also, we used a failure rate ≥5% to identify analytes for which testing may pose problems.

The analytes chosen for monitoring (**Tables 1 and 2**) are among the tests most often performed in physician's office and clinic laboratories, and they represent all major areas of clinical laboratory practice (chemistry, hematology, and microbiology). Failure rates were recorded from PT results submitted in 1994 (the first year that participation in a PT program was mandated by CLIA '88). Failure rates were subsequently re-examined in 1995, 2001, and 2004 for chemistry and hematology analytes (**Table 1**), and in 1995 to 1999, 2001, and 2004 for microbiology analytes (**Table 2**).

Table 1_Proficiency Testing Failure Rates for Chemistry and Hematology Analytes, 1994–2004[1]

Analyte	1994	1995	2001	2004
Cholesterol	18.7%	18.5%	4.2%	3.2%
Sodium	16.9%	17.6%	9.0%	5.5%
HDL cholesterol	16.4%	10.4%	10.2%	3.6%
Glucose	15.6%	21.1%	7.1%	2.4%
Prothrombin time	12.1%	7.0%	5.4%	3.2%
Potassium	6.3%	3.3%	0.6%	1.1%
Creatinine	5.7%	14.5%	6.5%	2.4%
Hemoglobin	4.3%	1.5%	1.2%	1.2%

[1]Failure rate = (number of unacceptable responses / total number of responses) × 100

Table 2_Proficiency Testing Failure Rates for Microbiology Analytes, 1994–2004[1]

Analyte	1994	1995	1996	1997	1998	1999	2001	2004
Positive cultures								
Genital/GC	23.3%	26.5%	18.2%	15.6%	20.4%	18.0%	7.1%	6.0%
Throat	14.9%	9.5%	6.5%	7.0%	4.7%	4.7%	5.4%	2.8%
Urine	35.0%	45.5%	27.1%	19.2%	13.7%	12.2%	16.2%	7.3%
Negative cultures								
Genital/GC	7.8%	7.1%	5.8%	5.6%	4.6%	6.0%	0.9%	0.6%
Throat	9.0%	7.2%	3.8%	4.2%	3.5%	2.8%	2.4%	1.0%
Urine	6.1%	5.2%	3.7%	4.3%	3.0%	3.6%	1.3%	1.9%
Gram stains	n/a[†]	18.5%	11.3%	13.7%	13.1%	8.9%	6.1%	5.4%

[1]*Failure rate = (number of unacceptable responses / total number of responses) × 100*
[†]*Gram stains were not included in API's proficiency testing programs in 1994.*

Results

Failure rates for all chemistry and hematology analytes, except hemoglobin, were above 5% in 1994, and they remained above 5% in 1995 for all analytes except hemoglobin and potassium (**Table 1**). By 2001, failure rates had declined significantly, although failure rates for sodium, HDL cholesterol, glucose, prothrombin time, and creatinine were still above 5% (**Table 1**). The decline continued through 2004, when only the failure rate for sodium remained greater than 5% (**Table 1**).

In 1994, failure rates for positive microbiology cultures ranged from 14.9% to 35.0% (**Table 2**). In 1995, failure rates for positive genital/GC cultures and positive urine cultures increased to 26.5% and 45.5%, respectively. Failure rates for negative cultures were also above 5% during 1994 to 1995. Although failure rates declined during the 10-year period, in 2004 they remained above 5% for both positive genital/GC cultures and positive urine cultures.

The failure rate for Gram stains was 18.5% in 1995 (the first year Gram stains were offered in API's proficiency testing programs) and declined to 5.4% in 2004.

Discussion

In general, the data suggests that, for most analytes, PT performance has improved in laboratories that were newly regulated by CLIA '88, but problems remain in microbiology.

Chemistry and hematology analytes. The high failure rates in 1994, and subsequent declines in failure rates, confirm the results of other studies that document declining failure rates as laboratories gain experience in proficiency testing.[1-3] From this result, we infer that testing of patient samples has likely improved as well. This inference is consistent with the conclusions of other researchers that improved PT performance positively impacts patient test performance.[2,4-5] One reason for this is that PT reveals systematic errors which can be corrected either by improved calibration and QC methods or by improved instrument design.[6-9] A second reason is that evaluating PT results provides opportunities to monitor staff competency and provide education, which can help improve performance.[6,9]

The exception to the overall positive trend with chemistry and hematology tests is that the failure rate for sodium remained above 5% throughout the 10-year period (**Table 1**), which may indicate a problem with the grading criteria for this analyte. A similar analysis of PT failure rates in Wisconsin found that failure rates for sodium ranged from 3% to 5% in 2002.[2] It would be useful to examine sodium failure rates from other PT providers to clarify whether testing sodium is problematic, and, if so, whether the problem is in testing the sample or in grading the result. If the latter, perhaps the CMS should consider implementing alternative grading criteria for sodium (which is currently the target value +/- 4 mmol/L) similar to the alternative grading criteria now used for glucose (the greater of the target value +/- 6 mg/dL or the target value +/- 10%).

Microbiology analytes. As was the case with chemistry and hematology analytes, PT performance with microbiology analytes improved from 1994 to 2004; however, the failure rates for positive genital/GC cultures, positive urine cultures, and Gram stains remained above 5% throughout the 10-year period (**Table 2**). Moreover, the failure rates for the positive cultures greatly exceeded the failure rates for the negative cultures. These results imply that many positive cultures and Gram stains from patient samples may be missed as well.

Laboratory performance by CMS criteria. The goal of our analysis was to assess PT performance by analyte and identify problem tests. Accordingly, the statistics in **Table 1** and **Table 2** were derived from CMS criteria that define unacceptable results for a particular test. It is instructive, however, to also examine PT performance using CMS criteria for unsatisfactory performance (that is, a score of less than 80% acceptable results in a test event). This is depicted in **Table 3**, which shows the numbers and percentages of laboratories with unsatisfactory performance on at least 1 proficiency test event in 2004. This data includes laboratories with unsuccessful performance as defined by CMS (that is, unsatisfactory performance on 2 consecutive events, or 2 out of 3 events). These statistics seem to indicate that testing of all analytes was consistently reliable in approximately 95% of laboratories. We believe, however, that assessment based solely on CMS criteria for satisfactory or successful laboratory performance can obscure serious problems in both individual laboratories and the broad population of laboratories.

Under CMS criteria, a laboratory that consistently scores 80% on PT events is judged successful in proficiency testing. This implies that, at least theoretically, up to 20% of a laboratory's patient test results could be unreliable, and PT would not detect this problem. This also means that, across the population of laboratories, significant problems may go undetected.

A second problem occurs in the way some tests are coded by CMS. For example, CMS aggregates all bacteriology tests (cultures, Gram stains, and susceptibility studies) under 1 code, which means the statistic in **Table 3** depicting unsatisfactory performance in bacteriology (5.1% of laboratories) is somewhat

Table 3_Laboratories With Unsatisfactory Proficiency Test Performance on One or More Events, 2004[1]

Analyte	Number of Laboratories	Unsatisfactory Performance Number (%)
Bacteriology[†]	4,156	212 (5.1)
Cholesterol	2,925	100 (3.4)
Sodium	3,333	163 (4.9)
HDL cholesterol	2,674	107 (4.0)
Glucose	3,568	126 (3.5)
Prothrombin time	2,073	108 (5.2)
Potassium	3,492	65 (1.9)
Creatinine	3,224	100 (3.1)
Hemoglobin	6,931	231 (3.3)

[1] "Unsatisfactory proficiency test performance" means that a laboratory scored less than 80% acceptable results on an event, as defined by CMS criteria.
[†] Bacteriology includes results for cultures and Gram stains.

misleading. Although this statistic provides information about the number of laboratories that have difficulty with bacteriology tests, it says nothing about which tests pose problems. In contrast, an analysis of failure rates for individual tests using CMS criteria for unacceptable results clearly shows that positive genital/GC cultures, positive urine cultures, and Gram stains may pose problems.

Limitations. Interpretation of the results is subject to 3 limitations. First, the data are not stratified by type of laboratory (physician's office, clinic, and small [<100 beds] hospitals). Thus, although the results reflect the performance of small laboratories, they do not exclusively reflect the performance of laboratories newly-regulated by CLIA '88.

Second, the extent to which improvements in PT performance signify improvements in testing patient samples is uncertain. This is because PT focuses on the analytical phase of the testing process; hence, it cannot detect errors in the pre- or post-analytical phases. Thus, it is possible for a laboratory to perform well on PT surveys and still perform poorly with patient samples due to uncorrected errors in the pre- and post-analytical testing phases.

Third, the data does not demonstrate trends in individual laboratories. Rather, it reflects performance trends in a changing population of laboratories. Laboratories that consistently perform poorly may stop testing, either because CLIA '88 requires them to cease testing if they consistently fail PT challenges, or because they no longer believe it is feasible to continue testing. As a result, the PT process may gradually select for laboratories that are proficient in testing, and a decline in failure rates does not necessarily reflect an improvement in the remaining laboratories' performance.

Conclusion

One impetus for the implementation of CLIA '88 was the perception that testing of patient samples was unreliable in many physician's office and clinic laboratories. Consequently, one goal of CLIA '88 was to improve test reliability in these laboratories. The results of this survey are encouraging because they suggest that, for most analytes, this goal has been met.

However, the results of this survey also suggest that laboratories miss many positive microbiology cultures and Gram stains. This alone is cause for concern, but other problems exist in microbiology as well. The literature documents that many small laboratories fail to follow recommended practices in microbiology, and that these problems have persisted for many years.[10-17] We believe microbiology testing urgently needs improvement, and we urge medical directors, laboratory administrators, and microbiology supervisors to focus attention in this area.

Finally, we would welcome further studies to assess how CMS coding practices and performance criteria may impact detection of problems. As illustrated by our microbiology data, this is an area that needs further investigation. LM

Acknowledgement: We thank Danny Bjoraker for compiling statistics for this report.

1. Centers for Disease Control and Prevention. Clinical laboratory performance on proficiency testing samples—United States, 1994. MMWR March 08, 1996;45(09):193-196. Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/00040598.htm. Accessed on: August 16 2006.
2. Ehrmeyer S, Laessig R. Has compliance with CLIA requirements really improved quality in US clinical laboratories? *Clinica Chimica Acta.* 2004;346:37–43.
3. Stull TM, Hearn TL, Hancock JS, et al. Variation in proficiency testing performance by testing site. *JAMA.* 1998;279:463–467.
4. St. John TM, Lipman HB, Krolak JM, et al. Improvement in physician's office laboratory practices, 1989–1994. *Arch Pathol Lab Med.* 124:1066–1073.
5. Tholen D, Lawson NS, Cohen T, et al. Proficiency test performance and experience with College of American Pathologists' programs. *Arch Pathol Lab Med.* 1995;119:307–311.
6. Jenny RW, Jackson-Tarentino KY. Causes of unsatisfactory performance in proficiency testing. *Clin Chem.* 2000;46:89–99.
7. Hoeltge GA, Phillips MG, Styer PE, et al. Detection and correction of systematic laboratory problems by analysis of clustered proficiency testing failures. *Arch Pathol Lab Med.* 2005;129:186–189.
8. Cembrowski GS, Carey RN. Adding value to proficiency testing programs. [Editorial]. *Clin Chem.* 2000;46:7–8.
9. Johnson PR. The contribution of proficiency testing to improving laboratory performance and ensuring quality patient care. *CLMR.* 2004;18:335–341.
10. Edson DC, Dyke JW, Glick T, et al. Identification of *Escherichia coli* O157:H7 in a proficiency testing program. *LabMedicine.* 2005;36:229–231.
11. Boyce TG, Pemberton AG, Wells JG, et al. Screening for *Escherichia coli* O157:H7—a nationwide survey of clinical laboratories. *J Clin Microbiol.* 1995;33:3275–3277.
12. Edson DC, Glick T, Massey LD. Susceptibility testing practices for *Streptococcus pneumoniae:* Results of a proficiency testing survey of clinical laboratories. *Diag Micro Inf Dis.* 2006;55:225–230.
13. Doern GV, Brueggermann AB, Pfaller MA, et al. Assessment of laboratory performance with *Streptococcus pneumoniae* antimicrobial susceptibility testing in the United States: A report from the College of American Pathologists Microbiology Proficiency Survey Program. *Arch Pathol Lab Med.* 1999;123:285–289.
14. Jones RN. Method preferences and test accuracy of antimicrobial susceptibility testing: Updates from the College of American Pathologists microbiology surveys program (2000). *Arch Pathol Lab Med.* 2001;125:1285–1289.
15. Kiehlbauch JA, Hannett GE, Salfinger M, et al. Use of the National Committee for Clinical Laboratory Standards guidelines for disk diffusion susceptibility testing in New York State Laboratories. *J Clin Microbiol.* 2000;38:3341–3348.
16. Stevenson KB, Samore M, Barbera J, et al. Detection of antimicrobial resistance by small rural hospital microbiology laboratories: Comparison of survey responses with current NCCLS laboratory standards. *Diag Microbiol Infect Dis.* 2003;47:303–311.
17. Novak RW. Do proficiency testing participants learn from their mistakes? Experience from the EXCEL throat culture module. *Arch Pathol Lab Med.* 2002;126:147–149.