

NBSIR 73-197 *Revised*

Clinical Laboratory Performance Analysis Using Proficiency Test Statistics

Peter W. Finkel, John W. Rowen

Technical Analysis Division
Institute for Applied Technology
National Bureau of Standards
Washington, D. C. 20234

December 1973

Final Report
(Supersedes COM--73--11253)

Available NTIS - #COM74 10468 (55page);

Prepared for
Division of Health Evaluation
Office of the Assistant Secretary for Planning and Evaluation
Department of Health, Education, and Welfare
Washington, D. C. 20201



NBSIR 73-197 ERRATA SHEET

DELETE ALL PERCENTAGE SIGNS (%) FROM TABLES C1 and C2. THEY INCORRECTLY
APPEAR AFTER EACH M.R.P. AND STD. DEV.

NBSIR 73-197

**CLINICAL LABORATORY PERFORMANCE
ANALYSIS USING PROFICIENCY
TEST STATISTICS**

Peter W. Finkel, John W. Rowen

Technical Analysis Division
Institute for Applied Technology
National Bureau of Standards
Washington, D. C. 20234

December 1973

Final Report
(Supersedes COM-73-11253)

Prepared for
Division of Health Evaluation
Office of the Assistant Secretary for Planning and Evaluation
Department of Health, Education, and Welfare
Washington, D. C. 20201



U. S. DEPARTMENT OF COMMERCE, Frederick B. Dent, Secretary
NATIONAL BUREAU OF STANDARDS, Richard W. Roberts, Director

Executive Summary

This report summarizes the first part of a two-part study performed for the Division of Health Evaluation, Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education, and Welfare by the Technical Analysis Division (TAD) of the National Bureau of Standards. The study reviewed the first two years (July 1969 to October 1971) of a proficiency testing program for laboratories licensed under the Clinical Laboratory Improvement Act (CLIA) of 1967 on the basis of data collected under the regulatory authority of the Center for Disease Control (CDC).

The major study effort was associated with statistical analyses of the data to:

- (1) Evaluate the temporal effect of continued participation in the CDC proficiency testing program.
- (2) Describe the population of licensed (interstate) laboratories in terms of various profile descriptors.
- (3) Determine, through one-way analyses of variance, whether the mean differences associated with the various levels of each independent variable (excluding number of employees) were significantly different.
- (4) Contrast the relative performance of licensed and volunteer laboratories.
- (5) Determine if bias is introduced by pooling various analytical methods in calculating the reference laboratory sample target value.

Statistical analyses were performed on two sets of data supplied by CDC, the agency within the Public Health Service (PHS) responsible for administering the CLIA '67 program. The first set, covering 268 licensed clinical laboratories, contained the analytical results (raw data, not scores) of proficiency tests conducted in clinical chemistry and hematology. Also contained in this set were the results of laboratory testing for several hundred other clinical laboratories who participate in the same proficiency tests but who are not licensed under the CLIA program (volunteer and reference laboratories).

The second set of data contained profile information on 237 of the 268 licensed laboratories, including a record of the number of full-time employees, the highest academic achievement level of the working supervisors, and the total number of accreditation or licensure programs in which the laboratory was involved. Corresponding profile data were not available for the volunteer or reference laboratories.

The temporal effect was evaluated through pooling the data over real-time. In this manner, laboratories were grouped according to their absolute point on the learning or experience curve. Inference to the performance of individual laboratories was not considered, nor was it ever intended to be a part of this study. Also inherent in the analyses

was the assumption of methods consistency independent of the level of each constituent. The major findings of the analysis are as follows:

1. It was not possible to discern meaningful changes over the two-year period in the overall level of licensed or volunteer laboratory performance as measured by the mean-accuracy approach (average relative difference from target value established by CDC reference laboratories) for any of the 13 constituents analyzed.
2. The variability among the licensed laboratories as measured by the standard deviation has decreased over the two-year period under study. This is an indication that the group performance has become more consistent. The same remark is true for the volunteer laboratory group as well, where the improvement was often much more marked.
3. Of the four independent variables studied, including Accreditation, State Program Intensity, Supervisory Education, and Analytical Method/Laboratories, only the last one appeared to be significant. That is, the general level of capability was about the same whether laboratories were or were not involved in proficiency testing programs other than CDC; whether they were located in states which did or did not have proficiency testing programs of their own; or whether the working supervisor had a B.S., M.S., Ph.D., or M.D. degree. Only the Analytical Method used (or the difference among laboratories) appeared to be significant ($p < .001$).
4. The data do not appear to support an argument in favor of establishing reference group target values on a method-by-method basis. More data, collected through the use of a statistically designed experiment which would avoid confounding analytical methods with laboratories by having each laboratory use several methods, would be helpful in reaching more rigorous and definitive conclusions.

The Clinical Laboratory Improvement Act of 1967, administered by the Center for Disease Control, provides for a many-faceted interlocking regulatory program. This study has looked at only one aspect of the CLIA '67 program, namely proficiency testing. Not included in the study, but also of major importance, are such functions as on-site laboratory reviews, training, written critiques in the qualitative specialties, and method development. There are few indications of meaningful changes in the average level of performance as measured by proficiency testing results since the licensed laboratories entered the CLIA proficiency testing program. Consideration should be given to alternative sampling methods, such as reduced or skip-lot sampling, for those constituents which appear to present no analytical challenge to the licensed and volunteer laboratories. Greater emphasis might then be placed on those constituents which give the laboratories the most difficulty (cholesterol and creatinine, for example).

This report includes a review of the legislative intent behind the law (CLIA '67) and an annotated bibliography of pertinent literature in the field of laboratory proficiency testing for the years 1963 to 1971.

Contents

	Page
Executive Summary	ii
1. Introduction	1
2. Background	2
2.1. Proficiency Testing Programs	3
3. Study Methodology	5
3.1. Inputs to the Analysis	5
3.2. Method	5
4. Analytical Results	7
4.1. Temporal Effect	7
4.1.1. Interstate Laboratories	8
4.1.2. Volunteer Laboratories	9
4.2. Profile Information: Licensed Laboratories Only	9
4.2.1. Additional Program Involvement	9
4.2.2. Number of Full-Time Employees	9
4.2.3. Number of Supervisors per Laboratory	10
4.2.4. Supervisor Academic Achievement	10
4.2.5. Academic Achievement of Supervisors Within Laboratories	11
4.3. Analyses of Variance	12
4.3.1. Accreditation	12
4.3.2. State Program Intensity	12
4.3.3. Supervisory Education	13
4.3.4. Analytical Method/Laboratories	13
4.4. Method Bias in Determining Target Values	13

Contents (continued)

	Page
5. Summary and Conclusions	15
5.2. Conclusions	15
Appendix A. A Review of the Legislative History of CLIA '67	17
Appendix B. Abstracts	21
Appendix C. Temporal Effect of Continued Participation in the CDC Proficiency Testing Program	35
Appendix D. Measuring Laboratory Performance in the CDC Proficiency Testing Program	49
Appendix E. Mean Measures of Relative Performance	51

Clinical Laboratory Performance Analysis Using Proficiency Test Statistics

Peter W. Finkel
John W. Rowen

The proficiency testing aspects of the Clinical Laboratory Improvement Act of 1967 program was assessed. The overall ability of licensed or volunteer laboratories to accurately determine mean values for any of the 13 constituents was not significantly altered during the first two years of program operation. However, the variability of the laboratories has decreased over the two-year period. It appears that the program has increased the consistency of laboratory performance.

The general level of laboratory capability seemed to be independent of involvement in state-supported or voluntary proficiency testing programs other than the CDC program and of whether the working supervisor had a B.S., M.S., Ph.D., or M.D. degree. Choice of analytical method did significantly affect performance. Although insufficient evidence was available to make a definitive statement, the data do not appear to support arguments favoring establishment of method-dependent reference group target values.

Finally, it appears that consideration should be given to alternative sampling methods, such as reduced or skip-lot sampling, for those constituents which appear to present no analytical challenge to the licensed and volunteer laboratories. Greater emphasis might then be placed on those constituents which give the laboratories the most difficulty (cholesterol and creatinine, for example).

Key words: Accuracy; CLIA '67; clinical chemistry; hematology; microbiology; precision; proficiency testing.

1. Introduction

This report summarizes the first part of a two-part study of clinical laboratory performance in the United States. The study was conducted by the Technical Analysis Division (TAD) of the National Bureau of Standards for the Division of Health Evaluation, Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education, and Welfare.

The objectives of this study were to determine if the participation of interstate (and volunteer) laboratories in the CDC proficiency testing program:

- (1) Has resulted in significant changes in average laboratory performance as measured by the accuracy (distance from target

- value) of proficiency testing results for these laboratories over a two-year period.
- (2) Reveals classes of laboratories which are more/less accurate from the standpoint of proficiency test results.
 - (3) Offers insights, conclusions or recommendations pertaining to the frequency of testing, constituents tested and laboratory analysis procedures used.

The purpose of this report is to publish the results of statistical analyses performed on an extensive data base which was compiled and maintained by the Center for Disease Control (CDC), Public Health Service, Health Services and Mental Health Administration, Department of Health, Education, and Welfare. The data for these analyses were the results of proficiency tests administered by the Center for Disease Control, extending for a two-year period beginning in July 1969; and various profile descriptors of the laboratories licensed under CLIA '67.

2. Background

Public Law 90-174 (Partnership for Health Amendments, 1967), Section 5, was enacted to assure consistent performance of accurate laboratory procedures and services. The law also includes provisions relating to the application for and issuance of licenses, quality control, personnel standards setting, and proficiency testing. The legislation impacts on clinical laboratories engaged in interstate commerce. The law was enacted in 1967, but was not fully implemented until July 1, 1969. The Center for Disease Control serves as the regulatory agency for this legislation.

Clinical laboratories that meet eligibility requirements under the Clinical Laboratory Improvement Act (CLIA) legislation are licensed by "procedure or category of procedure" such as: microbiology, serology, bacteriology, glucose, hemoglobin, etc. After receiving their license they are thereafter subject to periodic proficiency testing in many of the licensed areas. These tests are used as one means for determining the competency of the laboratory staff and the adequacy and quality of facilities, equipment, reagents and laboratory procedures. Normally, samples for the proficiency testing program are prepared and mailed to the licensed laboratory every third month. The analytical results obtained by the laboratory are subsequently returned to CDC for scoring, evaluation, and critique where applicable. (Appendix D provides a brief description of the program.) When a laboratory's performance reflects lack of proficiency (significant deviation from norms established by the CDC reference laboratories and/or peer group analysis) corrective measures are taken by CDC. The extent of the corrective action varies according to the individual situation. In instances where poor performance occurs repeatedly, where cooperation is lacking or suggested remedial action is not taken by the laboratory, suspension or revocation of the license can result.

Appendix A of this report provides a brief summary of events leading to the enactment of CLIA '67, changes in the administration of the act which have occurred since its enactment, and details concerning the purpose and sub-objectives of the program as administered by CDC.

2.1. Proficiency Testing Programs

Proficiency testing programs are monitoring systems designed to provide a basis for ascertaining the reliability of laboratory performance. These programs serve to prompt the appropriate actions needed to assure results that are medically useful and interconvertible (the results from one method can be translated to comparable results from a different method). Proficiency testing often involves the analysis of submitted samples; however, it is important to remember that this is but one component of any total proficiency assurance program.

Some other objectives, or components, of a multifaceted program such as CLIA '67 are:

- a. To define the "state-of-the-art" in terms of:
 - (1) Methodology
 - (2) Reagents
 - (3) Standards
 - (4) Personnel performing test
- b. To define sources of error in terms of:
 - (1) Procedure
 - (2) Education and performance of analyst
 - (3) Reagent
 - (4) Equipment (includes improper use of equipment as well as use of improper equipment for an assay)
 - (5) Monitoring systems
- c. To define criteria to substantiate "medical usefulness"
- d. To define educational needs and stimulate remedial measures.

The Federal Regulations published on December 31, 1968, clearly describe CDC's regulatory authority with regard to unacceptable performance by a licensed laboratory in the proficiency testing program. The reader is referred to the published regulations for a complete description of this topic, but in brief, it is incumbent upon a licensed laboratory to maintain satisfactory performance in all proficiency testing programs in which it is involved. Failure to do so can result in revocation of licensure or exemption therefrom.

The extent to which the existing proficiency programs do, in reality, impact on general laboratory performance improvement is still not widely agreed upon. The difficulty in determining whether proficiency testing programs have improved performance is due to one or more of the following:

- a. Changes in methods of interpretation of the results.

- b. The effect of new laboratory entries in the program whose initial results obscure the improved changes of the established participating members.
- c. Technological changes in the laboratory testing.
- d. The fact that virtually all proficiency testing programs, including CLIA '67, require laboratories to analyze samples and specimens which are identifiable as testing material. The laboratory that knows it is being tested can use more experienced personnel, better quality control procedures and more accurate testing procedures for these sample analyses. In some cases, the licensed laboratory has split the sample received from CDC, analyzed one portion itself, and sent the other portion to another laboratory for confirmative analysis. Thus, proficiency testing can result in measures which are representative of the best work that a laboratory is capable of instead of its routine performance.

A review of the literature indicates that it has been extremely difficult to determine whether these programs have indeed accomplished the goal of improving laboratory performance. There is a paucity of data and much of what is available is conflicting. For example, proficiency test data¹ from the College of American Pathologists (CAP) indicates a significant improvement over the last 20 years. Proficiency testing data obtained by the New York City Health Department², however, indicates no significant improvement, albeit for a shorter period of time, over the years 1964-1968. Other published studies³ support the New York City findings by demonstrating the inability of a laboratory to perform equally well in two different proficiency testing programs.

The aforementioned findings and speculations should be considered when proficiency testing is suggested as anything other than one of several indices to be used for laboratory evaluation. Perhaps a more germane question to ask is whether proficiency testing in combination with other measurement and control procedures has any utility.

¹L. P. Skendzel, D. J. Hanson, and W. H. Civen, "The 1949 College of American Pathologists Survey Revisited," American Journal of Clinical Pathology, LIV (1970), 493.

²M. Schaeffer, D. Wedelock, P. S. May, S. Blatt, and M. Wilson, "The Clinical Laboratory Improvement Program in New York City II, Progress After Five Years of Experience," Health Laboratory Science, VII (1970), 242.

³R. G. Martinek, et al., "Comparison of Methods for Conducting Laboratory Proficiency Surveys," Health Laboratory Science, V (1968), 239.

3. Study Methodology

3.1. Inputs to the Analysis

The analysis was performed on data compiled and maintained by the Center for Disease Control. The data were of two types:

- a. Profile information for most of the licensed (interstate) laboratories, including types of accreditation or licensure programs in which active, state in which located, number of full-time employees, number and academic training of working supervisors.
- b. Proficiency test results for licensed, voluntary, and reference laboratories covering the period of July 1969 through September 1971. These data pertain to the following constituents of human blood: glucose, creatinine, urea nitrogen, uric acid, total protein, cholesterol, calcium, sodium, potassium, chloride, hematocrit, red blood cell count (RBC) and hemoglobin. The analytical method used by the laboratory is also reported. If a laboratory is licensed in a category (e.g., clinical chemistry) or subcategory (e.g., cholesterol) for which CDC has developed a proficiency testing program, it receives a shipment of test samples every three months. The general procedure calls for three samples per shipment, usually one at each of the high, medium and low level of each constituent.

In addition to the technical data, information was obtained on the "strength" of legislation and controls within states in which interstate and volunteer laboratories are located. An extensive literature search (Appendix B) was also conducted to gather information for the development of a suitable analysis design and for the interpretation of results.

3.2. Method

The methods used for analysis of the data were relatively simple, with the major activities devoted to re-formatting to detect discernible temporal effects, editing, and augmentation of basic data already compiled on magnetic tape. Computer programs were developed to group and statistically analyze this information as follows:

- a. The raw data (analytical test results) were separated on a constituent-by-constituent (test-type) basis.
- b. A target value was calculated for each sample, using the reference laboratory analyses. In each instance where there were at least 30 reference laboratories reporting results on a method-by-method basis, the target value consisted of the average analysis for those 30 or more reference laboratories. When there were fewer than 30 reference laboratories using one particular method, the data were pooled and one target value was calculated for the methods involved. That one target value was then used for all of the methods which had been pooled.

- c. The difference between the analytical result reported by a laboratory in the test and the target value calculated from the reference laboratory data was used as the measure of performance. The difference was normalized by dividing it by the target value. For example, a licensed laboratory analyzed a sample of cholesterol by method A and reported 185 mg/100 ml as its result. The average reference laboratory analysis (calculated from 63 different laboratories all using method A) was 166.5 mg/100 ml. The difference ($185 - 166.5 = 18.5$) was divided by the target value; $18.5/166.5 = 0.11$, and the result was the "normalized" measure of performance for that laboratory for that sample.
- d. Each laboratory's data file was searched in order to determine when it first entered the CDC proficiency testing program. This entry date was designated "time zero" for that laboratory, establishing its position in a pseudo-time scale.
- e. The normalized difference (performance) data for interstate laboratories were grouped according to the pseudo-time scale, and group averages were calculated. That is, all of the differences representing "first" analyses of cholesterol were pooled together (by method of analysis) to give an average difference for "time zero." The same procedure was followed for each subsequent time period. This data is summarized in Table C-1, Appendix C.
- f. The volunteer laboratory data were treated exactly as was described in the preceding paragraph, with one exception. Whereas the first entry into the data bank for almost all of the licensed laboratories actually represented their first sample analyses, there are a small number of licensed laboratories which were participating in the CDC proficiency testing program prior to the enactment of CLIA '67. Hence, their "first" analyses were at a much earlier time, and these records are not part of the CLIA '67 proficiency testing data base. Consequently, even though the volunteer laboratory data covers the same time frame as the interstate data, it does not represent the same segment of the overall learning or experience curve. This data is summarized in Table C-2, Appendix C.

Target values established by CDC reference laboratories were used as the standard of comparison in this study. No analysis has been done, however, to test the validity of these target values. It might be informative to point out some characteristics of the reference laboratory group. For example, a laboratory which is licensed in one or more categories may serve as a reference (or referee for the qualitative specialties) laboratory in areas in which it is not licensed. Many of the CDC reference laboratories rank among the leading state health departments and teaching hospitals. There do not appear to be any prescribed ground rules for selecting reference laboratories except that they must be willing to analyze the samples and report the results back to the originating source.

The computer programs used in this study provided the following outputs:

1. A temporal analysis of laboratory proficiency for all participating laboratories for each constituent test method. If laboratories were "improving" over time, mean differences from the estimated target values should have become smaller. If no "improvement" occurred, convergence would not have been noted.
2. An analysis of variance program determining whether performance (accuracy with which laboratories approximated target values) is dependent upon laboratory characteristics such as size, accreditation, and/or academic training of the supervisor.

These outputs were analyzed and are presented in this report, together with conclusions upon which to base future action on the part of the Department of Health, Education, and Welfare.

4. Analytical Results

4.1. Temporal Effect

Participation in a proficiency testing program may result in changes in individual laboratory performance. Changes in overall group performance are not as easily discerned. However, there are at least two ways in which changes among the laboratories might be detected. An analysis of the general level of accuracy which clinical laboratories achieve as a consequence of continued participation in a proficiency testing program would provide one indication of change. The "scatter" (standard deviation) of results obtained among the participating laboratories would be a second indication. As the group of laboratories becomes more homogeneous the between-laboratory standard deviation will be smaller.

Several assumptions are inherent in the statistical interpretation of the data. The following assumptions are representative:

1. The distribution of the relative differences between a laboratory's result and the target value is not functionally related to the analytical method used, the level of the constituent, or the constituent itself, except for a possible learning effect which may vary from one constituent to another.
2. The normalizing technique employed, whereby the observed difference between a laboratory's result and the target value is divided by the target value $[(O-E) \div E]$, satisfactorily negates any effect which the level of the constituent might have had on the difference.
3. The groups of laboratories (the size of which increased by a factor of at least three during the two years under study) are equally representative at each point in time. That is, those laboratories which entered the interstate licensure program early in its development do not vary significantly in their general level of performance from laboratories which entered the program at a later date. This assumption also implies that early or late entrance into the program is not associated with the analytical methods used.

Paragraphs 4.1.1 and 4.1.2 discuss measures of performance for both the interstate and volunteer laboratories.

4.1.1. Interstate Laboratories

This phase of the statistical analysis examined the question of whether participation in CDC's proficiency testing program results in a changed level of performance over time for laboratories in general. The question was examined on the basis of overall accuracy at various time points after entry; i.e., by the mean performance among participating laboratories and by the consistency among the participating laboratories, at various time points after entry into the program. Improved performance would be reflected by smaller mean differences and by smaller standard deviations, respectively, with continued program participation. A linear regression analysis was performed on these values, for each constituent, over the two-year period 1969-1971.

The constituents for which quantitative data was available and for which the regression analyses were performed were: glucose, creatinine, urea nitrogen, uric acid, total protein, cholesterol, triglycerides, calcium, sodium, potassium, chloride, hematocrit, red blood cell count, and hemoglobin.

On the basis of nine or less quarters of testing, it was not possible to identify significant shifts in mean differences for any constituents at the five percent level of probability.

The analysis also identified a phenomenon which has not been satisfactorily explained. Sometimes there was a consistent (but not statistically significant) bias between the mean differences for licensed laboratories and reference laboratory target values. For example, Table C-1 (Appendix C) shows that when creatinine was analyzed by the alkaline picrate (w/o Lloyds) method, the target value (the average analysis for all the reference laboratories which used that same method) was always larger than the mean computed for the interstate group using the same method. This was also true for uric acid when analyzed by the phosphotungstate carbonate method and cholesterol analyzed by the AutoAnalyzer and SMA methods. Positive biases (i.e., the target value was consistently smaller) were not as prevalent; they were observed only in the following instances: (1) urea nitrogen, all methods; (2) red blood cell count, all methods; (3) hematocrit, specifically the micro Hematocrit method, as well as for all methods; and (4) chloride. No definitive information has been found to explain these biases.

The variability of the interstate laboratories, as demonstrated by the pooled standard deviations, indicates that the group is becoming more homogeneous. From Table C-1, Appendix C, it is evident that most of the standard deviations become smaller as time in the proficiency testing program increases. Exceptions to this trend are found for creatinine and red blood cell count.

4.1.2. Volunteer Laboratories

Volunteer laboratories were, on the average, performing at slightly lower proficiency levels for some constituents in 1971 than they were in July 1969, (i.e., cholesterol and red blood cell count). Obvious exceptions to this are urea nitrogen (all methods), uric acid (all methods), and hemoglobin (all methods), where the average differences are small.

As with the interstate group, the volunteer laboratories are generally becoming more homogeneous. In fact, for seven constituents, the standard deviation for volunteer laboratories was smaller after eight or nine quarters than the comparable interstate statistic. Considering that the volunteer laboratory standard deviation was initially larger for 10 out of the 13 constituents, the group as a whole has made greater strides towards increased homogeneity than have the interstate laboratories.

4.2. Profile Information: Licensed Laboratories Only

The licensed laboratories were characterized according to several attributes reported to CDC by the laboratories. For instance, it was possible to determine those licensed laboratories which are involved in one or more additional accreditation or licensure programs, the number of employees, supervisor's educational achievements, and other characteristics. Paragraphs 4.2.1 through 4.2.5 discuss the various characteristics of the 237 licensed laboratories for which data was available.

4.2.1. Additional Program Involvement

Of the 237 licensed laboratories for which data was available, 12 percent (29) do not participate in any other inspection, accreditation or licensure programs. The 208 laboratories that do participate in other programs are distributed as follows:

	<u>Number</u>	<u>Percent</u>
Participate in one additional program	68	29
Participate in two additional programs	93	39
Participate in three additional programs	47	20
	<hr/>	<hr/>
Subtotal	208	88
Do not participate in other programs	29	12
	<hr/>	<hr/>
Total	237	100%

4.2.2. Number of Full-Time Employees

More than half of the 237 laboratories have less than eight full-time employees; 18 percent, or 42 laboratories, have more than 20 employees. The frequency distribution below describes this characteristic:

	<u>Number</u>	<u>Percent</u>
Laboratories with 3 or less employees	53	19
Laboratories with 4-7 employees	75	35
Laboratories with 8-11 employees	33	14
Laboratories with 12-20 employees	34	14
Laboratories with more than 20 employees	42	18
Total	<u>237</u>	<u>100%</u>

4.2.3. Number of Supervisors per Laboratory

Six percent of the licensed laboratories do not have any supervisors, while 71 percent of the laboratories have 1-4 supervisors. The breakdown is as follows:

	<u>Number</u>	<u>Percent</u>
Laboratories with no supervisors	15	6
Laboratories with 1 supervisor	100	43
Laboratories with 2 supervisors	46	19
Laboratories with 3 or 4 supervisors	46	19
Laboratories with 5-12 supervisors	24	10
Laboratories with more than 12 supervisors	6	3
Total	<u>237</u>	<u>100%</u>

4.2.4. Supervisor Academic Achievement

There were 620 individuals identified as working supervisors in the 237 laboratories. Of these 620 supervisors, roughly 17 percent have no college degree. The following table contains the overall frequency of supervisors having attained various levels of academic achievement.

	<u>Number</u>	<u>Percent</u>
Supervisors with no college degree	106	17
Supervisors with a B.S. or A.B. Degree	259	42
Supervisors with an M.S. or M.A. Degree	53	9
Supervisors with a Ph.D. Degree	72	11
Supervisors with an M.D. Degree	80	13
Supervisors with other degrees	50	8
Total	<u>620</u>	<u>100%</u>

4.2.5. Academic Achievement of Supervisors Within Laboratories

A cross-tabulation by academic achievement and number of supervisors per laboratory was also prepared; that is, how the academic degrees are distributed among laboratories with different numbers of supervisors. The breakdowns were obtained for laboratories with one, two, and three supervisors. These tabulations follow:

Laboratories with One Supervisor

	<u>Number</u>	<u>Percent</u>
Supervisors with no college degree	19	19
Supervisors with a B.S. or A.B. Degree	46	46
Supervisors with an M.S. or M.A. Degree	9	9
Supervisors with a Ph.D. Degree	5	5
Supervisors with an M.D. Degree	14	14
Supervisors with other degrees	7	7
	<hr/>	<hr/>
Total	100	100%

Laboratories with Two Supervisors

<u>First Supervisor</u>	<u>Second Supervisor</u>	<u>Frequency</u>
No college degree	No college degree	6
No college degree	B.S. or A.B.	9
No college degree	M.D.	1
No college degree	Other	1
B.S. or A.B.	B.S. or A.B.	9
B.S. or A.B.	M.S. or M.A.	2
B.S. or A.B.	Ph.D.	4
B.S. or A.B.	Other	1
M.S. or M.A.	Ph.D.	1
M.S. or M.A.	M.D.	1
M.S. or M.A.	Other	1
Ph.D.	Ph.D.	1
Ph.D.	M.D.	2
Ph.D.	Other	1
M.D.	M.D.	5
Other	Other	2
		<hr/>
		47

Laboratories with Three Supervisors

<u>First Supervisor</u>	<u>Second Supervisor</u>	<u>Third Supervisor</u>	<u>Frequency</u>
None	B.S. or A.B.	M.S. or M.A.	2
None	B.S. or A.B.	M.D.	1
None	B.S. or A.B.	Other	2

Laboratories with Three Supervisors (continued)

<u>First Supervisor</u>	<u>Second Supervisor</u>	<u>Third Supervisor</u>	<u>Frequency</u>
None	None	Other	1
B.S. or A.B.	B.S. or A.B.	None	1
B.S. or A.B.	B.S. or A.B.	B.S. or A.B.	3
B.S. or A.B.	B.S. or A.B.	M.D.	1
B.S. or A.B.	B.S. or A.B.	Other	3
M.S. or M.A.	Ph.D.	M.D.	1
M.S. or M.A.	M.S. or M.A.	B.S. or A.B.	1
M.S. or M.A.	M.S. or M.A.	Ph.D.	1
Ph.D.	Ph.D.	Ph.D.	1
Ph.D.	Ph.D.	M.S. or M.A.	1
M.D.	M.D.	M.D.	3
M.D.	M.D.	B.S. or A.B.	1
M.D.	M.D.	Ph.D.	3
			26

All other combinations had a zero frequency; i.e., in none of the licensed laboratories that had three supervisors were all three without any college degree.

4.3. Analyses of Variance

A one-way analysis of variance of differences among test results associated with each of the independent variables (excluding number of employees) was performed for every constituent, using the NBS OMNITAB II computer program. Specifically, the values of relative differences for all laboratories and all time periods were classified according to one of the independent variables, and the differences among the means for this classification were investigated. The results of these analyses are summarized in Paragraphs 4.3.1 through 4.3.4.

4.3.1. Accreditation

The extent to which interstate laboratories participated in other accreditation or licensure programs did not significantly (tested at the .05 level of probability) affect their average performance in the CDC proficiency testing program. That is, the average difference for laboratories not involved with any other program was not significantly different from a comparable value associated with laboratories involved with several other programs. This finding was consistent for all 14 constituents.

4.3.2. State Program Intensity

Analysis of variance indicates that performance in the CDC proficiency testing program is not functionally related to intensity of programs at the state level, for all 14 constituents.

4.3.3. Supervisory Education

Different levels of education achieved by the working supervisors did not effect the performance of licensed laboratories in the proficiency testing program. For all constituents, there was no indication that laboratories whose work is supervised by persons with Ph.D. degrees performed at different levels than their counterparts whose supervisors had baccalaureate degrees or no degree at all.

4.3.4. Analytical Method/Laboratories

The data indicate that interstate laboratory performance in the CDC proficiency testing program is significantly ($p < .001$) affected by the analytical method used. This finding was consistent for all 14 constituents. However, the laboratories included in the study were not assigned at random to the various methods, and each laboratory used one and only one method to analyze each constituent. Therefore, unless one is willing to make the necessary assumption that interstate laboratories are a homogeneous group, and that the differences observed are truly due to the analytical methods alone and not to factors already analyzed, there are no firm conclusions which can be drawn from this simple one-way analysis of variance. Additional analyses of variance, which consider the possible interactions of the four independent variables, will be conducted in the near future and reported separately. Appendix E consists of summary tables developed from this one-way analysis of variance.

4.4. Method Bias in Determining Target Values

An original premise underlying this entire evaluative effort was that one measure of a laboratory's analytical performance is its ability to come close to the "right answer." In this study the "right answer," for each sample, was considered to be the average analytical value reported by a group of at least 30 reference laboratories, all using the same analytical method as the licensed laboratory. The limited number of cases where at least 30 laboratories used the same method restricted the number of method comparisons which could be made. Thus, the available data base was insufficient to conclusively analyze the importance of establishing target values on a method-by-method basis. Nevertheless, the limited results were suggestive. Table 4.1 summarizes the number of times the method comparisons were possible for each constituent.

The tests of significance, which were performed on the means (t test with unequal sample size, known variability) of the groups, did not indicate any statistical significance, at the 10 percent level, between the methods compared.

Table 4.1. Summary of Number of Times Method Comparisons Were Possible for Each Constituent

Constituents	Analytical Methods	Number of Method Comparisons
Glucose	AutoAnalyzer <u>vs</u> SMA	3
Creatinine	Alkaline Picrate (w/o Lloyds) <u>vs</u> AutoAnalyzer	5
Urea Nitrogen	AutoAnalyzer <u>vs</u> SMA	6
Uric Acid	AutoAnalyzer <u>vs</u> SMA	6
Uric Acid	Phosphotungstate (carbonate) <u>vs</u> AutoAnalyzer	3
Uric Acid	Phosphotungstate (carbonate) <u>vs</u> SMA	3
Total Protein	Biuret (serum) <u>vs</u> Refractive index	3
Total Protein	Biuret (serum) <u>vs</u> AutoAnalyzer	3
Total Protein	Biuret (serum) <u>vs</u> SMA	3
Total Protein	Refractive index <u>vs</u> AutoAnalyzer	3
Total Protein	Refractive index <u>vs</u> SMA	3
Cholesterol	Direct serum (acetic anhydride) <u>vs</u> AutoAnalyzer	3
Cholesterol	Direct serum (acetic anhydride) <u>vs</u> SMA	3
Cholesterol	AutoAnalyzer <u>vs</u> SMA	3
Calcium	Atomic absorption <u>vs</u> AutoAnalyzer	3
Calcium	Atomic absorption <u>vs</u> SMA	3
Calcium	AutoAnalyzer <u>vs</u> SMA	3
Sodium	Flame photometry <u>vs</u> AutoAnalyzer	8
Sodium	Flame photometry <u>vs</u> SMA	3
Potassium	Flame photometry <u>vs</u> AutoAnalyzer	11
Potassium	Flame photometry <u>vs</u> SMA	3
Potassium	AutoAnalyzer <u>vs</u> SMA	3
Chloride	Cotlove titrator <u>vs</u> AutoAnalyzer	3
Chloride	Cotlove titrator <u>vs</u> SMA	3
Chloride	AutoAnalyzer <u>vs</u> SMA	3
Hemoglobin	Cyanmethemoglobin (manual) <u>vs</u> Coulter S	3

5. Summary and Conclusions

5.1. Summary

This report summarizes the first part of a two-part study conducted for the Division of Health Evaluation, Office of the Assistant Secretary for Planning and Evaluation, Department of Health, Education, and Welfare. The study consisted of an examination of data from the chemistry and hematology portions of the first two years of a proficiency testing program for laboratories licensed under the Clinical Laboratory Improvement Act of 1967. Also included in the report is Appendix A, a review of the legislative history behind the law, and Appendix B, an annotated bibliography of pertinent literature in the field of laboratory proficiency testing from 1963 to 1971.

Statistical analyses were performed on two sets of data supplied by the Center for Disease Control. One set, covering 268 licensed clinical laboratories and several hundred volunteer laboratories, included the results of proficiency tests conducted in these blood constituents: glucose, creatinine, urea nitrogen, uric acid, total protein, cholesterol, triglyceride, calcium, sodium, potassium, chloride, hematocrit, red blood cell count and hemoglobin. The second set of data contained certain profile information on 237 of the 268 licensed laboratories. This information included the number of full-time employees, the highest academic achievement level of the working supervisors, a measure of the strength of the proficiency testing program operated by the state in which the laboratory was located, and the total number of accreditation or licensure programs in which the laboratory was involved.

5.2. Conclusions

It is very important to keep the following three aspects of proficiency testing programs in general, and the CDC program in particular, in mind when reviewing the data and formulating conclusions.

1. Almost every proficiency testing program measures the capability of participating laboratories instead of their routine level of performance. The extent to which capability and performance are related is unknown.
2. Factors such as scoring methods, criteria for determining outliers, and inter-convertibility among methods are known to affect measured performance in proficiency testing programs. Greater standardization among the major testing programs would be desirable.
3. Proficiency testing is only one aspect of a broad based quality assurance program. It is also necessary to have well trained laboratory personnel, fresh reagents, standard solutions of high quality, and timely feedback to the laboratory.

The following conclusions fall into two major categories. The first three deal specifically with the data analyses. The fourth conclusion

is identified with the proficiency testing program currently being conducted by the Center for Disease Control.

1. The only independent variable for which the mean differences were significantly different was Analytical Method/Laboratories. The other three variables, Accreditation, State Program Intensity, and Supervisory Education, were apparently unimportant.
2. The standard deviation, a measure of the variability among laboratories, has generally decreased over the two-year period under study for both the interstate and volunteer laboratory groups. This is an indication that both groups have become more homogeneous in their testing of proficiency program samples.
3. The data do not support establishing reference group target values on a method-by-method basis. More data, collected in a statistically designed experiment, would be helpful in reaching more rigorous and definitive conclusions.
4. There are few indications of meaningful changes in the average level of performance of laboratories participating in the CLIA proficiency testing program. Consideration should be given to a sampling scheme somewhat akin to skip-lot sampling for those constituents which appear to present no analytical challenge to laboratories. Greater emphasis might then be placed on those constituents which give licensed laboratories the most difficulty (cholesterol and creatinine for example).

Appendix A. A Review of the Legislative History of CLIA '67

Introduction

A review of the pertinent legislative material, including House and Senate report hearings and Federal Register (FR) announcements was made in search of answers to questions such as:

1. At the time of the hearings, what was the projected magnitude of the licensing effort? Approximately how many clinical laboratories were engaged in interstate commerce?
2. What were the ostensible problems in the clinical laboratory industry which led to the hearings and ultimate legislation?

Public Law and Rule Making

The Honorable Harley O. Staggers, Chairman of the House Committee on Interstate and Foreign Commerce, introduced HR 6418 early in 1967, at the request of the Administration. A similar version, S 724, was introduced in the Senate, and the bill which was eventually enacted was entitled "Clinical Laboratory Improvement Act of 1967, Section 353 of the Public Health Service Act." The intent of the Administration was to extend and strengthen the Partnership for Health Provision, Public Health Service Act, PL 89-749, by placing primary reliance upon state health authorities and their services programs. Broadly stated, the administration viewed the partnership as one where the Federal government would assist the states with financial and technical resources without assuming any authority for the conduct of the state programs.

Three major witnesses for the administration were Undersecretary of HEW, Wilbur J. Cohen; Surgeon General, Dr. William H. Stewart; and Communicable Disease Center Director, Dr. David J. Sencer. Dr. Sencer testified that 25 percent of all clinical laboratory analyses were so gross as to be medically useless. This testimony, from so eminent a witness, made a significant impact upon the congressional committee before which Dr. Sencer was appearing.

The major mileposts with respect to publishing the new legislation and its projected impact on the industry were:

1. October 15, 1968—A notice of rule making was published in the Federal Register (33 FR 15297-15303) which proposed the addition of a new Part 74 to the Public Health Service Regulations.
2. December 31, 1968—A notice that the new Part 74 was added to the PHS Regulations, effective 30 days after publication (January 31, 1969). This notice described the nine major subparts of the new Part 74 as follows:
 - Subpart A - Definitions and Applicability
 - Subpart B - License: Application and Issuance
 - Subpart C - Quality Control

Subpart D - Personnel Standards
Subpart E - Proficiency Testing
Subpart F - Accreditation—State Loans—Stringency of Standards—Termination
Subpart G - General Provisions
Subpart H - Revocation, Suspension, and Limitation of Licenses and Letters of Exemption Notice
Subpart I - Hearings (Reserved)

3. April 5, 1969—Effective date of the Act postponed again to July 1, 1969 (it had already been postponed from January 31 to April 1, 1969 in paragraph 2, Subpart I of FR No. 253).
4. July 1, 1969—The College of American Pathologist proficiency testing program was determined to be equal to or more stringent than the Federal program. Laboratories enrolled in the CAP program were granted exemption from CLIA upon submitting an application for same.
5. August 15, 1969—Effective date for cytotechnologists to become qualified extended from January 1, 1960 to January 1, 1969.
6. October 14, 1970—Notice in the FR that DHEW proposed to amend the CLIA '67 regulations with regard to granting hearings for affected laboratories.
7. October 25, 1969—HEW published a notice of intent in the FR to require a letter of exemption from laboratories accepting fewer than 100 specimens per year in a specific category or subcategory.
8. December 3, 1969—A notice published in the FR to amend CDC's Statement of Organization, Functions, and Delegation of Authority.
9. February 7, 1970—A list of laboratories which were granted exemption from CLIA '67 was published by HEW in the FR.
10. April 16, 1970—HEW published a notice in the FR that item 7 above was now in effect; exempt laboratories would have to have a letter of exemption, issued by the director, CDC.

Licensure—Eligibility and Exemption

There has never been a complete census of all operating clinical laboratories in the United States. Such an undertaking would be facilitated by a consensus as to a working definition of "a clinical laboratory." Current and past estimates of the total number of laboratories range from 15,000 to perhaps 200,000; depending upon whether physicians' offices are included.

When the director of the Center for Disease Control (CDC) testified at the House hearings, he estimated that "Ten percent of the clinical laboratories in this country are in interstate commerce." There are other references to the number of laboratories in interstate commerce, such as that which appears on page 2083 of the Congressional and Administrative News, 90th Congress (First Session, 1967): "Estimates given to the Committee by the Department of Health, Education and Welfare indicate that there are approximately 1,000 laboratories substantially operating in interstate commerce. These are primarily independent laboratories and a few hospital laboratories." The Administration's intent, in

supporting this legislation, was to impact on the largest segment of the industry as evidenced by the following quotation from the same 90th Session Congressional and Administrative News: "The committee believes that the Secretary of Health, Education, and Welfare should be required to provide an exemption from licensing in those cases where laboratory operations in commerce are so small or infrequent as not to constitute a significant threat to the public health."

There were many forces acting on the Congress during its consideration of the legislation, such as the life insurance industry which wanted exemption for laboratories whose primary business was associated with performing tests in conjunction with physical examinations for policy holders, the medical community which wanted any physician directed laboratory to be exempt from federal intervention, and various professional (society) accreditation organizations which operate financially successful quality control programs. Consequently, when Public Law 90-174, "The Partnership for Health Amendments of 1967" was enacted on December 5, 1967, it contained the following exemptions on page 580 of the U.S. Code—

(2) A license issued under this section shall be valid for a period of three years, or such shorter period as the Secretary may establish for any clinical laboratory or any class or classes thereof; and may be renewed in such manner as the Secretary may prescribe. The provisions of this section requiring licensing shall not apply to a clinical laboratory in a hospital accredited by the Joint Commission on the Accreditation of Hospitals or by the American Osteopathic Association, or a laboratory which has been inspected and accredited by such commission or association, by the Commission on Inspection and Accreditation of the College of American Pathologists, or by any other national accreditation body approved for the purpose by the Secretary, but only if the standards applied by such Commission, association, or other body in determining whether or not to accredit such hospital or laboratory are equal to or more stringent than the provisions of this section and the rules and regulations issued under this section, and only if there is adequate provision for assuring that such standards continue to be met by such hospital or laboratory; provided that any such laboratory shall be treated as a licensed laboratory for all other purposes of this section.

and on page 582—

(i) The provisions of this section shall not apply to any clinical laboratory operated by a licensed physician, osteopath, dentist, or podiatrist, or group thereof, who performs or perform laboratory tests or procedures, personally or through his or their employees, solely as an adjunct to the treatment of his or their own patients; nor shall such provisions apply to any laboratory with respect to tests or other procedures made by it for any person engaged in the business of insurance if made solely for purposes of determining

whether to write an insurance contract or of determining eligibility or continued eligibility for payments thereunder.

(1) Where a State has enacted or hereafter enacts laws relating to matters covered by this section, which provide for standards equal to or more stringent than the provisions of this section or than the rules and regulations issued under this section, the Secretary may exempt clinical laboratories in that State from compliance with this section.

As of January 18, 1972, there were 552 clinical laboratories licensed by the Center for Disease Control out of approximately 1,500 which are "probably" eligible from the standpoint of being engaged in interstate commerce. The 63 percent not licensed are assumed to be exempt because they accept less than 100 samples per year in interstate commerce or are otherwise accredited or licensed.

Summary

The committee hearings and related evidence indicated that the performance of clinical laboratories was not what it could or should be. The committee recognized that improvement of the clinical laboratories and their performance was a long run effort for the state governments, their health agencies and their public health service authorities. However, the committee also recognized that it would take time for the state licensing programs to be developed throughout the nation. In the meantime, the state programs would vary considerably with respect to procedures and quality. Because of this variance, laboratories doing business in interstate commerce could locate in states with weaker licensing programs or in states without programs. In addition to the need for licensing interstate business of the clinical laboratories, there was the above-mentioned administration and congressional plan to extend and strengthen the partnership for health provision of the PHS Act. These elements, the need for regulations of interstate samples and the extension of the partnership for health, were the motivating factors for CLIA '67. However, an unforeseen development in certain states has, to some extent, bypassed the anticipated effect of the law. At least two states have passed laws which prohibit solicitation of laboratory work from laboratories outside that state.

Appendix B. Abstracts

The entries in this annotated bibliography on proficiency testing programs in the open literature have been obtained from a search of the Department of Commerce National Technical Information Center; the Defense Documentation Center; the Index Medicus of the National Library of Medicine, DHEW library; from references in the papers themselves, and from a report entitled "Clinical Laboratory Evaluation Programs" dated July 31, 1970, by the Auerbach Corporation, Philadelphia. This last report was procured by the Department of Health, Education and Welfare on Contract 110-69-434.

The 34 entries on proficiency programs are arranged by years starting with 1963, and are alphabetical within years.

1963

1. Tonks, D. B., "A Study of the Accuracy and Precision of Clinical Chemistry Determinations in 170 Canadian Laboratories," Clinical Chemistry, 9:217 (1963).

The data obtained have shown that there was a lack of both accuracy and precision in many laboratories. There is obviously a need for an improvement in the methods of detecting errors and of evaluating precision in the laboratory. An adequate quality control program and the proper usage of control serums are indispensable for these purposes. It is also useful for every laboratory to take part in periodic studies such as this one in order to obtain an unbiased evaluation of its performance.

1967

1. Copeland, B. F., W. J. Blake, R. J. Muelling, and L. P. Skendzel, "A Report of the Standards Committee of the College of American Pathologists," American Journal of Clinical Pathology, 48:104 (1967).

The purpose of this survey was to stimulate interest in the accuracy and precision of chemical measurements commonly made in laboratories for the diagnosis and therapy of disease. The following measurements were included in the survey: glucose, urea nitrogen, bilirubin, cholesterol, calcium, phosphorous, sodium, potassium, bicarbonate, uric acid, and chloride.

2. Kaufmann, William and R. Vanderlinde, "Medical Laboratory Evaluation," New England Journal of Medicine, 277:1024 (1967).

The overwhelming evidence points to rather serious deficiencies in the performance of tests, in their reproducibility and thus in their accuracy and dependability as aids in clinical diagnosis. A built-in fallacy in assessing laboratory performance is the assumption that the test specimen is treated in the same fashion as the daily routine specimen

1967 (continued)

whereas this, in fact, rarely happens. The data show that low quality in testing is related to a low annual number of tests performed. The fact that there is a serious laboratory problem in this country should be recognized.

3. Schaeffer, M., D. Widdock, S. Blatt, and M. E. Wilson, "The Clinical Laboratory Improvement Program in New York City: I. Methods of Evaluation and Results of Performance Tests," Health Laboratory Science, 4:72 (1967).

This report is limited to the data obtained in evaluating the capabilities of all the laboratories within the jurisdiction of New York City and to factors that relate to the quality of performance.

In October 1966 there were 359 clinical laboratories with permits to operate in New York City. A different group of laboratories was tested each week, and each laboratory was tested at least four times in a two-year period. The tests performed included microbiology, syphilis serology, cross-matching and hemoglobin determinations. The standard deviation of all the results of the reference laboratories was calculated, the acceptable range being two standard deviations from the mean. About 50% of the independent laboratories and about 40% of all of the hospital laboratories were rated as unsatisfactory in bacteriology, while about 25-30% were rated as excellent. In clinical chemistry, about 50% were fair, and the remainder equally divided between unsatisfactory and excellent. The poorest results appear among the independent laboratories, although the other laboratories were not much better off. About 18% of the cross-matching tests were unsatisfactory; in syphilis serology 84% of the results of independent laboratories and more than 90% of the results of hospital laboratories were excellent.

Laboratory performance did not appear to be related to whether the director was full-time or not. Qualifications of the director did not appear to matter. The organization of the laboratory did appear to influence performance. Higher test accuracy was noted in the laboratories that processed high volumes of specimens.

There appeared to be a direct relationship between the use of professionally trained supervisors and the quality of performance. Forty excellent laboratories, representing only 29.6% of voluntary hospitals, employed 63.3% of all the supervisors in voluntary hospitals. This level of supervision is difficult to obtain and too expensive for a small laboratory. This paper expresses the opinion that the existing situation (1964-1966) in clinical laboratories cannot be permitted to continue indefinitely.

1. Barnett, Roy N., "Medical Significance of Laboratory Results," American Journal of Clinical Pathology, 50:671 (1968).

Tests performed in clinical laboratories are among the main diagnostic aids available to physicians. The great demand for these tests and the enormous number performed have led to Federal specification through the Medicare Act and the Clinical Laboratory Improvement Act of 1967 (CLIA '67). Regulations under both Acts set basic standards of operation and require participation in a proficiency testing program.

The basic question is how accurate must clinical laboratory work be? The Standards Committee of the College of American Pathologists has been the major proficiency surveying body in the United States for a number of years, and these surveys lead to two questions: (1) How well are laboratories performing certain common analyses? and (2) Is this kind of performance adequate for good medical practice? This provisional report examines these questions. Six (sic) terms are defined, including accuracy, precision, standard deviation, coefficient of variation, percentiles, normal range, medically significant, state-of-the-art, and decision level. Other topics include guidelines in establishing limits for reporting values of medical significance, a table of medically significant values, interlaboratory and intralaboratory evaluation, and the utilization of specific limits in medical practice.

2. College of American Pathologists, "Guidelines for Evaluating Laboratory Performance in Survey and Proficiency Testing Programs," American Journal of Clinical Pathology, 49:457 (1968).

Yearly surveys have produced data that describe the accuracy and reproducibility of procedures. Variability increases as variables are introduced. Single laboratory precision is better than multiple laboratory precision.

In consultation with Dr. W. J. Youden, NBS, the CAP established criteria for acceptability in interlaboratory surveys, but not for precision objectives by individual laboratories. Reference laboratory results will be provided to survey participants. Performance terminology is given, as are good performance criteria for hemoglobin, glucose, urea nitrogen, calcium, phosphorous, total protein, chloride, uric acid, albumin, bilirubin, sodium, potassium, creatinine, and cholesterol.

3. Martinek, R. G., et al., "Comparison of Methods for Conducting Laboratory Proficiency Surveys," Health Laboratory Science, 5:239-256 (1968).

The program reported here was planned and carried out by the Illinois State Department of Public Health. The survey method used, whereby unknowns are analyzed by participating laboratories, is reported and summarized in the report. Another method relied on the use of patients' data (PD), wherein each laboratory was instructed to list 90 consecutively

1968 (continued)

determined values. These distributions were plotted, and normal curves fitted, each being described by the mean or average and the corresponding standard deviation. Normal ranges were established for each test used. From the data from the conventional surveys by the American Association of Blood Banks and CAP correlation analyses were made as well as studies of within and between surveys using Kendall's coefficient of rank correlation. Results "...indicated as close a relationship between a proficiency sample and the patients' means as between proficiency samples from the two agencies." Considerable clarification of interpretation was obtained. "When the conventional proficiency survey results are interpreted in this manner the conclusion is inescapable that the patients' data is informative and directly related to our concern." The results of testing patients' specimens represents everything that the laboratory does. The analyses of patients' data yield standard deviations that are similar in magnitude to those from conventional and extensive surveys of laboratories for prothrombin time. Conclusions from these analyses are (1) do not use either technique as the basis for clinical laboratory regulatory action at this time, and (2) carefully evaluate the various quality control methods.

4. Prier, J. E., L. Sideman and I. J. Yankevitch, "Clinical Laboratory Proficiency Testing," Health Laboratory Science, 5:12 (1968).

Clear evidence documents deficiencies in hospital and independent clinical laboratories (CL). Available methodology provides ways to obtain high accuracy and consistency if properly applied. CL evaluation involves three factors: (1) personnel qualifications, (2) on-site inspection, and (3) proficiency testing. Proficiency testing (PT) in Pennsylvania was made mandatory in 1961. Two responses to law are noted: (a) well trained pathologists provide the least objections; (b) "notable" objections are from directors who are marginal in both training and facilities. PT must be carried out to assure competency of CL work. It is doubtful that an adequate CL control program can be obtained by voluntary participation. (Gives history of PT in Pennsylvania.) Any effective program must be enforced. Ten characteristics of a new PT program are discussed. Competency in CL work must include PT, and PT without enforcement is inadequate, collecting only disquieting data. A PT program must reach all laboratories equally. There should be no provision for waiver because of education of director or other reasons. Accurate laboratory work is related to close supervision. Inspections and "on-the-spot" testing are essential parts of a total evaluation program. Programs provided by legislation with provisions for compulsory PT are mandatory.

5. Report from the Ad Hoc Committee on "Documentation of Procedures, Diagnostic Criteria, and Results in Laboratory Medicine of the Committee on Pathology," Division of Medical Science NRC-NAS-NAE, American Journal of Clinical Pathology, 49:757-760 (1968).

Conclusions: (1) That there is a need for documentation of medical laboratory data in the United States in order to promote the comparability

1968 (continued)

of data over long periods and between different sections of the country as patient mobility increases.

(2) That the questionnaire method has been successful in determining the state-of-the-art as practiced in leading medical institutions.

(3) That information on the state-of-the-art should continue to be collected and should be published.

(4) That a Conference on Documentation of Procedures, Diagnostic Criteria, and Results in Laboratory Medicine should be held.

(5) That there appears to be a definite need for a Center for Documentation of Procedures, Diagnostic Criteria, and Results in Laboratory Medicine. The obvious benefit to patient care makes this an important goal.

1969

1. Eilers, R. J., "Total Quality Control in the Medical Laboratory," Hospital Progress, 50:92 (1969).

The concept of total quality control for the medical laboratory offers a means of achieving the professional objective of improving the quality of the service that it renders. Quality Control does not mean best, but it could. The word "control" in the phrase Quality Control includes a four-part management tool:

- Establishment of standards
- Determination of conformance to these standards
- Taking corrective action
- Planning for improved standards.

Nine benefits from improved Quality Control in laboratories include:

- Improvement in quality of laboratory results
- Improvement of employee morale
- Improvement of inspection methods
- Establishment of firm time standards
- Establishment of preventive maintenance schedules
- Development of a factual basis for cost accounting.

The cost of obtaining and maintaining a level of quality in a laboratory is consolidated with the cost from failure to obtain the given level of quality. Four elements of cost are cited: (1) Prevention cost, (2) Appraisal costs, (3) Costs of internal failure, and (4) Costs of external failures. A total Quality Control program can be divided into three phases:

- Assign quality responsibility to all employees
- Employ a quality control specialist
- Develop special techniques for all laboratory functions.

Six phases of Quality Control are:

- New design control
- Incoming material control

Process control
Production output control
Product reliability
Special process studies.

The College of American Pathologists has an inspection and accreditation activity for total laboratory services that is accepted by the JCAH and CDC. This service has been in operation for many years. Participants in 1968 numbered 3900.

To obtain high Quality Control, all in the chain must recognize the existing deficiencies and initiate programs for raising the existing level, otherwise responsibility for this function—Quality Control—will become the responsibility of state or Federal bureaucrats.

2. Hendry, I. A., Editor, "Reports of Clinical Chemistry Surveys," Board of Education of the College of Pathologists of Australia, Sydney meeting, August 1969.

The program reported here was based on the premise that most of the testing done in clinical laboratories is on specimens from people in relatively good health. The test plan was designed to provide information (a) about the level of laboratory daily precision, and (b) on the ability of the laboratories to maintain precision over a three-month interval.

Two classes of data were collected, the one from a conventional survey where two sets of different sera were mailed for testing, and the other consisting of two groups of data from tests on patients' specimens, each group containing the results of 90 tests. The sets and groups consisted of data taken in the same two time intervals. The patients' test data were the same as for the conventional survey, glucose, urea, total protein, and cholesterol. Unknown to the laboratories, only two different concentrations were used, with at least one specimen of each in any given mailing of the six. Conclusions cited are that:

- (a) The specimen concentrations in the human sera remained constant throughout the three months of the study.
- (b) Information regarding the accuracy and precision of individual laboratories can be obtained from patients' test data.
- (c) Surveys of laboratories consisting of a few survey unknowns are of limited value.
- (d) Surveys are useful for determining the performance of laboratories as groups, but are of practically no value for estimating the performance of individual laboratories.
- (e) The information from patients' test data is at least as informative as that obtained from tests performed on survey samples.
- (f) Patients' test data is sufficiently detailed to be reasonably valid for estimating individual laboratory performance.

1969 (continued)

This report concluded by recommending that the College of Pathologists of Australia establish a computing service for laboratories, with feedback to individual laboratories.

3. Hoffman, Robert G., "Patients' Test for Quality Control," Clinical Chemistry, 15:533 (1969).

A brief outline of the method of "average-of-normals" is given. It consists in the plotting of the mean \pm the 95% confidence limits as a function of the tests or time. Each point was the average of 16 tests in one demonstration, and of 10 in the other. In this way it is stated that out-of-normal shifts can be seen.

4. Kanon, Dor, "Computer Analyses Lab Tests, Does Own Quality Control," Modern Hospital, 113:105 (1969).

In a form-phase, all-day operation, this chemistry laboratory computer updates patient files, standardizes laboratory values, does quality control, and produces cumulative reports. The clinical chemistry laboratory at St. Vincents Hospital, New York, handles 285,000 tests a year. On-line operation has form phases that describe in detail the day's work. The use of standards is described, as is the reliance on quality control programs. Loss of sensitivity from reagent deterioration, and increase in standard deviation from dialyzer diaphragm deterioration have been detected. Monthly cost is about \$7,000. Biochemistry production increased 51% in the first six months of operation with no increase in personnel.

5. Kaufmann, William and six co-authors, "Clinical Laboratory Performance Experience with the New York State Clinical Laboratory Program," New York State Journal of Medicine, 15:1989 (1969).

In 1964 the New York State Legislature passed the Laboratory Improvement Act, which requires all clinical laboratories to operate under permit and to have in charge an individual who holds a valid certificate of qualification. There were 430 clinical laboratories included in the 1967 to 1968 survey, made up of 229 hospital laboratories and 201 independent laboratories, privately or municipally owned. The six areas of bacteriology, mycobacteriology, mycology, serology, hematology, and clinical chemistry were covered.

6. McSwiney, R. R. and D. A. Woodrow, "Types of Errors Within a Clinical Laboratory," Journal of Medical Laboratory Technology, 26:340 (1969).

In this study all errors were counted over four weeks covering a total of approximately 7,000 tests. There were 163 errors found, 63 of them by senior technicians and 57 by pathologists. In addition, two are known to have "escaped" into the wards. "Organizational" errors consisted

mostly of omitted or delayed analyses, and accounted for 34% of the total. Over half the errors (54%) were clerical in origin. A normal day's work produced 15 to 20 errors. Errors are made more than realized; they are forgotten as a familiar part of life but are not permissible within the clinical laboratory.

7. Sideman, L., D. T. Wilson, and J. J. Murphy, Jr., "A Survey of Immunohematological Proficiency Among 131 Clinical Laboratories in Pennsylvania," Health Laboratory Science, 6:156-161 (1969).

Survey objectives were: (1) to uncover laboratory errors in the identification of major blood groups and Rh type, and (2) having found the errors, to offer consultation and training in order to eliminate the causes. The overall improvement shown in a subsequent re-evaluation of these same laboratories emphasizes the value of a comprehensive evaluation program which includes interlaboratory studies, consultative services, training, and re-evaluation.

8. Seracci, R., "Factors Affecting Accuracy and Precision in Clinical Chemistry," American Journal of Clinical Pathology, 52:161-166 (1969).

Accuracy and precision in four clinical chemistry determinations (glucose, urea, total cholesterol, total proteins) have been investigated in a survey of 404 Italian laboratories. Inaccuracy and imprecision were found which appear, under survey conditions, to be little related to factors such as class and size of laboratories, qualifications and seniority of analyst, and method of analysis. This would imply that acting upon these factors is less important in improving laboratory performance than is the systematic use of quality control methods.

9. Skendzel, L. P. and W. J. Youden, "A Graphic Display of Interlaboratory Test Results," American Journal of Clinical Pathology, 51:161 (1969).

A simple graphic system worked out by Youden is described and advantages noted. One requirement for this system is that laboratories analyze two samples for the same constituent. Reference laboratory results are plotted against each other, with intersecting lines passing through each mean, the SD is multiplied by 3.035 to obtain the radius of a circle that is drawn with the intersection of the means as the center. This circle will then include 99% of the values of reports for the two samples. Two tangents with positive slope at 45° are drawn to the circle one at each end of the diameter. Pairs of values from other participants are then used to plot additional points.

1. Barnett, Roy N. and W. J. Youden, "A Revised Scheme for the Comparison of Quantitative Methods," American Journal of Clinical Pathology, 54:454-462 (1970).

The authors, on the basis of experience gained in evaluation of 50 different "kits" for chemical analysis, have developed a revised and simplified method for conducting such evaluations. The scheme, which is applicable for any comparison of quantitative methods, describes reproducibility studies (20 once-a-day analyses), recovery studies in triplicate at two or three levels, and comparison studies with patient serum, all performed by the test method and a reference method. Statistical calculations and criteria of acceptability are described in detail and a complete example given.

2. Barnett, Roy N., W. Harold Civin, and Irwin Schoen, "Multiphasic Screening by Laboratory Tests - An Overview of the Problem," American Journal of Clinical Pathology, 54:483-492 (1970).

The authors considered the problems of multiphasic screening by laboratory tests in the framework of the concept "Total Quality Control in the Clinical Laboratory." After reviewing the data, the problems, and the opinions of others, they drew the following conclusions. It is entirely premature to recommend the routine use of extensive multiple screening tests for either hospital admission or general populations, considering the present ignorance of physicians about every one of the six categories we used. At present these multiple screening tests should be considered research rather than service activities. Although the laboratory problems in testing are still formidable, they are being solved far more rapidly than are the problems relating to the medical usefulness of the test results. There is an urgent need for approximately controlled, large scale, multidisciplinary studies to answer the basic questions concerning the utility of the data. We cannot accept the assumption that the production of huge volumes of "screening" information will by itself contribute to human knowledge or health. A selected bibliography is appended.

3. Eilers, Russell J., "Total Quality Control for the Medical Laboratory: The Role of the College of American Pathologists Survey Program," American Journal of Clinical Pathology, 54:435 (1970).

Six phases of Quality Control are set forth as follows:

- I. Selection of proper test method
- II. Selection of proper standards control and SRM's
- III. Internal Quality Control of process and proficiency testing
- IV. Proper format of results
- V. Reliability and interpretation of results
- VI. Inspection and accreditation.

1970 (continued)

4. Ivenson, M. A., "The Need for Analytical Chemistry in Clinical Chemistry Training Programs," Analytical Chemistry, 42:53A (1970).

The shortage of adequately trained clinical chemists will soon become a manpower crisis, but recent Federal fund tightening activities have dampened enthusiasm for new programs. There are 1,600 persons in the U.S. who have more than a bachelor's degree in chemistry and are members of the American Association of Clinical Chemists. The 20 most common clinical chemistry tests are listed. Each clinical chemistry laboratory doubles its own size every five years. A short history of clinical is given, starting with Paracelsus in 1526. Comments on the administrative and academic arrangements usually formed by clinical chemists are made. Need for analytical chemical excellence in clinical chemistry laboratories is stressed.

5. Cavan, Thomas L. and John W. King, "An Evaluation of the Microbiology Portions of the 1969 Basic, Comprehensive, and Special College of American Pathologists Proficiency Testing Surveys," American Journal of Clinical Pathologists, 54:514 (1970).

The proficiency testing program of the CAP was markedly expanded in 1969 by adding three new surveys: Special Bacteriology, Special Mycobacteriology, and Special Mycology. The use of mycobacterial and mycologic specimens in the Comprehensive and Basic Surveys was expanded.

The proficiency testing surveys of the CAP and those of other agencies are limited to evaluating a laboratory's ability to identify a given organism or organisms. No satisfactory methods exist for large scale surveys in microbiology, based on simulated clinical specimens, to enable laboratory proficiency evaluation in isolation techniques as well as the identification of bacteria and fungi.

In all, 83 specimens of various bacteria and fungi were distributed, with 19 being common to two or more surveys. A total of 34,908 reports from participants have been evaluated. Methods of rating participant performance in terms of referee performance are given, and the results of both are presented in seven tables. In general, 80% or more of the participant laboratories were able to identify the pyogenic bacteria submitted in a good or acceptable manner. The exceptions are cited and discussed. Mycobacteria were acceptably identified by significantly fewer laboratories than the pyogenic bacteria. Fungi were more poorly identified than pyogenic bacteria, but were correctly identified more frequently than the mycobacteria.

The quality and scope of survey material have been improved, as well as the results reported by participants. Improved techniques for the conduct of such surveys are given, including specimen checking and in freeze drying and handling specimens.

1970 (continued)

6. Gilbert, Roy K., "Analysis of Results of the 1969 Comprehensive Chemistry Survey of the College of American Pathologists," American Journal of Clinical Pathology, 54:463-482 (1970).

The samples of this survey consisted of eight vials mailed out in pairs on four occasions during the year. The results are described for analyses for glucose, calcium, total bilirubin, potassium, sodium, urea nitrogen, uric acid, creatinine, chloride, phosphorous, protein-bound iodine, and total protein. The total number of analyses was 140,730. The results were categorized by method used, and a clear pattern of relationship was apparent between the method and both the mean values and the standard deviations. Some methods showed a constant relationship in mean values with one method higher or lower throughout the series. In others there was variability and the relationship changed. This fluctuation could occasionally be traced to interaction with other constituents. The methods also showed a relationship to the standard deviations throughout the series, with one method usually higher or lower than another. The relationship tended to be more constant than that for the mean values.

Variance is ascribed to three sources: the basic analytical performance, calibration standards, and systematic differences between laboratories; these are discussed at length. The results show that many methods give mean values that differ from those of other methods. The data also show that this is not a constant relationship and that interaction between constituents occurs. The spread of values is wider by some methods than others, and the automated method tends to produce a narrower range although a survey offers some bias in favor of automated means.

7. Hanson, Daniel J., "Suggested Definitions for Clinical Laboratory Standards and Reference Materials," American Journal of Clinical Pathology, 54:451 (1970).

The CAP is concerned about the existing confusion regarding definition and nomenclature of standards that are applicable to clinical pathology. The situation is an impediment to laboratory performance evaluation and to the application of standards and controls in the clinical laboratory.

There is an impelling and immediate need for agreement on terms and definitions by pathologists, technologists, and manufacturers to resolve confusion and permit evaluation of reports. Definitions are presented and discussed for Primary Standard, Primary Volumetric Standard, and Secondary Volumetric Standard. The clinical versions of these standards are also presented, as is the term Reference Material or Reference Control, originally referred to by Radin as "reference samples."

1970 (continued)

8. Koepke, John A., "Immunohematology Proficiency Testing, 1966-1969," American Journal of Clinical Pathology, 54:508-511 (1970).

Immunohematology proficiency testing as measured by the Laboratory Surveys of the College of American Pathologists, from 1966 to 1969, is reviewed. During this period there has been a gradual expansion of this portion of the survey. Over the years studied there has been an evident improvement in ABO grouping and the detection of atypical antibodies. Crossmatching has continued to be well done; however, performance of Rh typing has been spotty, presumably due to variation in typing sera or red cell testing specimens, or both.

9. Reed, A. H., "Use of Patient Data for Quality Control of Clinical Laboratory Tests," Clinical Chemistry, 16:129 (1970).

The average-of-normals quality-control method is examined mathematically. Two modifications are proposed. After modification, the essential unchanged feature of the method is the idea of censoring the laboratory test results so that only means of test values inside the normal range are plotted on the quality-control chart. Several limitations of the modified method are discussed, and it is concluded that, although censoring does improve the probability of detecting drift, more statistical research and further modifications are needed before the average-of-normals method will be useful for quality control in the clinical laboratory.

10. Rodwell, V. W., Editor, "Test Theory, Proficiency Tests, and Unevaluated Tests," Test of the Month No.1, Proficiency Testing Service, American Association of Bioanalysts (1970).

Research is needed to gain a greater understanding of procedure and objectives in carrying out proficiency testing and on-site inspections. PT under Medicare is cited as is CLIA '67. Cost problems of private groups are noted. CAP participant cost is \$881/year, while CDC will do the same thing for less than \$125.

There is little understanding of the theory of testing. It was first given formal expression by Dr. Reuben L. Kahn: first, a procedure is employed to derive a probability statement with regard to any determination; second, group validity in regard to any determination is always more reliable than individual validity; third, the parameters involved in any test are sensitivity, specificity, and reproducibility of results. An exception is blood typing where there is no room for error in determining ABO and Rh types.

1970 (continued)

11. Schaeffer, M., D. Widelock, P. S. May, S. Blatt, and M. E. Wilson, "The Clinical Laboratory Improvement Program in New York City: II. Progress After Five Years of Experience," Health Laboratory Science, 7:242 (1970).

In previous studies, clinical laboratories in New York City were analyzed with respect to organization and technical competence. From that evaluation it became apparent that laboratories performing satisfactorily are those that employ full-time competent supervisors. This experience led to efforts to improve the performance of consistently poor laboratories.

Two methods designed to improve the proficiency of clinical laboratories in either bacteriology or chemistry were evaluated. On-site consultation by experts in these two disciplines, and workshop instruction were studied for their effect on results of proficiency testing of laboratories selected for assistance. The latter were compared with similar laboratories treated routinely, without advice or instruction.

Analysis of results appears to indicate that some improvement in performance was achieved simply through repeated proficiency tests. Laboratories participating in bacteriology workshops showed improved performance as compared with control laboratories. However, laboratories enrolled in clinical chemistry workshops showed no greater improvement than control laboratories.

12. Schoen, I., "Guidelines to Laboratory Management," California State Department of Public Health, Some Proficiency Testing Experiences in California for 1968, Volume D-1 (1970).

Proficiency testing is mandatory for many licensed laboratories in California. There were 226 laboratories in the State Comprehensive Survey, and 201 in the Basic Survey; the Basic Survey includes sufficient specimens and examinations to meet the mandatory PT requirements, while the comprehensive survey includes more specimens and examinations than required by the mandatory regulations. Overall, the performance of California laboratories and all participating laboratories was very similar with no notable difference in the data. The means of reporting and selection of specimens by current survey techniques appears to lead to unsatisfactory erroneous mixed-up reports involving up to 1.5% of reports, and up to 5% of laboratories.

Automation seems more precise for some procedures. This precision might be due not to automation itself, but may be more appropriately due to standardizing or controlling instrumentation methods, reagents, and standards. A "correct" report on a survey specimen does not mean a "correct" interpretation of a report on a patient's specimen from a laboratory with an incorrect set of normal values. Medical Usefulness Criteria as well as data from Peers or Reference Laboratories should be utilized in evaluating performance.

1970 (continued)

Local or regional data accumulation and evaluation should be considered to achieve more prompt evaluation with follow-up testing and consultation for unsatisfactory performance.

To best serve the interest of reliable, better, and improved laboratory services by a specific laboratory for the patient, a program of standards for quality control analogous to standards for a Blood Transfusion Service should be considered.

13. Skendzel, Laurence P., "Guidelines for the Design of Laboratory Surveys," American Journal of Clinical Pathology, 54:437-447 (1970).

This report outlines the system used in the College of American Pathologists National Survey Program for medical laboratories with special emphasis on data handling and evaluation of a laboratory's performance. The statistical techniques for data analysis of quantitative procedures include: (1) exclusion of widely deviant values, (2) calculation of mean, SD, and CV after exclusion of deviant values to describe the distribution of values, (3) calculation of the average CV to summarize the distribution of values from several surveys, (4) reliance on the two-sample plot and the "t" test to combine results derived by different methods. Evaluation of laboratory performance is based on the "State-of-the-Art" concept which uses participant's values to set acceptance and not acceptable limits. Other systems for evaluation are discussed.

Appendix C. Temporal Effect of Continued Participation
in the CDC Proficiency Testing Program

Table C1. Temporal Effect of Interstate Laboratory Participation in
CDC Proficiency Testing Program

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>GLUCOSE</u>										
AutoAnalyzer	M.R.P.* (%)	0.018	-0.041	0.004	0.007	-0.009	-0.027	-----	-----	-----
	Std. Dev. (%)	0.1135	0.1190	0.0561	0.0686	0.0637	0.0660	-----	-----	-----
	N**	100	83	75	64	57	45	31	-----	-----
SMA	M.R.P. (%)	0.006	0.002	-0.023	0.032	0.010	-----	-----	-----	-----
	Std. Dev. (%)	0.0790	0.0795	0.1426	0.0643	0.0662	-----	-----	-----	-----
	N	100	82	63	52	37	-----	-----	-----	-----
All Methods	M.R.P. (%)	0.012	-0.004	-0.032	0.011	0.001	-0.018	-0.029	0.002	-----
	Std. Dev. (%)	0.2517	0.1194	0.1350	0.0880	0.0944	0.1022	0.0780	0.0451	-----
	N	372	276	218	166	132	100	68	31	-----
<u>CREATININE</u>										
Alkaline Picrate, w/o Lloyds	M.R.P. (%)	-0.013	-0.053	-0.079	-0.068	-0.113	-0.067	-0.076	-----	-----
	Std. Dev. (%)	0.5641	0.2013	0.2493	0.1985	0.2271	0.1772	0.1620	-----	-----
	N	175	156	130	115	93	78	49	-----	-----
All Methods	M.R.P. (%)	-0.030	-0.061	-0.088	-0.079	-0.101	-0.065	-0.072	-0.032	-----
	Std. Dev. (%)	0.4932	0.2082	0.2579	0.1993	0.2401	0.1692	0.1612	0.7589	-----
	N	333	245	190	164	136	107	65	34	-----

* M.R.P. = Mean Relative Performance

**N = Number

Table C1. Temporal Effect of Interstate Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)											
	0	3	6	9	12	15	18	21	24			
<u>UREA NITROGEN</u>												
AutoAnalyzer	M.R.P. (%)	-.007	-.016	-.005	-.015	-.004	-.014	-----	-----	-----	-----	-----
	Std. Dev. (%)	0.0921	0.0887	0.1611	0.0747	0.1284	0.0685	-----	-----	-----	-----	-----
	N	92	77	65	53	44	39	-----	-----	-----	-----	-----
SMA	M.R.P. (%)	-.019	0.029	-.036	0.024	0.001	0.010	0.011	-----	-----	-----	-----
	Std. Dev. (%)	0.1218	0.1092	0.1025	0.1319	0.0633	0.1484	0.0657	-----	-----	-----	-----
	N	107	89	66	60	48	40	32	-----	-----	-----	-----
All Methods	M.R.P. (%)	0.054	0.021	0.026	0.014	0.022	0.024	0.008	0.010	-----	-----	-----
	Std. Dev. (%)	0.2551	0.1926	0.2848	0.2409	0.1760	0.1389	0.1220	0.0766	-----	-----	-----
	N	355	262	204	167	140	111	75	40	-----	-----	-----
<u>URIC ACID</u>												
Phosphotungstate Carbonate	M.R.P. (%)	-.019	-.029	-.055	-.036	-.038	-.018	-----	-----	-----	-----	-----
	Std. Dev. (%)	0.2796	0.2433	0.1600	0.1652	0.1642	0.1507	-----	-----	-----	-----	-----
	N	86	66	56	50	39	31	-----	-----	-----	-----	-----
AutoAnalyzer	M.R.P. (%)	0.023	0.020	0.022	0.003	-.051	-----	-----	-----	-----	-----	-----
	Std. Dev. (%)	0.1514	0.1438	0.1673	0.2886	0.1752	-----	-----	-----	-----	-----	-----
	N	74	65	58	52	37	-----	-----	-----	-----	-----	-----
SMA	M.R.P. (%)	-.0340	-.0088	-.009	-.013	-.033	-.040	-.029	-----	-----	-----	-----
	Std. Dev. (%)	0.1815	0.1844	0.1834	0.2659	0.1780	0.3071	0.1798	-----	-----	-----	-----
	N	103	88	66	56	46	40	31	-----	-----	-----	-----

Table C1. Temporal Effect of Interstate Laboratory Participation in CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>URIC ACID</u>										
All	M.R.P. (%)	- .013	- .006	- .014	- .014	- .041	- .010	- .051		
	Std. Dev. (%)	0.2207	0.1967	0.1783	0.2445	0.1745	0.2124	0.1732	0.022	
	N	338	259	210	175	138	113	72	44	
<u>CHOLESTEROL</u>										
AutoAnalyzer	M.R.P. (%)	- .014	- .069	- .034	- .062	- .027	- .051			
	Std. Dev. (%)	0.1151	0.1484	0.1107	0.1590	0.1392	0.1367			
	N	77	56	49	42	39	33			
SMA	M.R.P. (%)	- .009	- .031	- .041	- .064	- .012	- .06			
	Std. Dev. (%)	0.1425	0.1377	0.1424	0.1348	0.1002	0.1593			
	N	108	81	62	53	44	31			
All	M.R.P. (%)	- .009	- .028	- .034	- .050	- .029	- .048	- .026	- .045	
	Std. Dev. (%)	0.1428	0.1457	0.1400	0.1482	0.1215	0.1405	0.1333	0.1228	
	N	436	299	224	184	148	112	81	44	
<u>TOTAL PROTEIN</u>										
Biuret, serum std.	M.R.P. (%)	.012	- .003	.002	.015	- .015				
	Std. Dev. (%)	0.0673	0.0701	0.0646	0.0772	0.0528				
	N	101	69	57	42	35				

Table C1. Temporal Effect of Interstate Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>TOTAL PROTEIN (continued)</u>										
SMA	M.R.P. (%)	- .010	- .007	0.006	0.013	- .005	- .009	- .007	- .010	-----
	Std. Dev. (%)	0.0637	0.0514	0.0565	0.0682	0.0521	0.0339	0.0373	0.0373	-----
	N	128	98	84	75	66	52	47	38	-----
All	M.R.P. (%)	- .006	- .003	0.005	0.016	- .007	- .007	- .006	- .013	- .006
	Std. Dev. (%)	0.0677	0.0626	0.0627	0.0735	0.0531	0.0430	0.0400	0.0375	0.0306
	N	384	276	220	182	151	110	88	58	36
<u>SODIUM</u>										
Flame Photometry	M.R.P. (%)	0.002	0.006	0.006	0.007	- .003	- .003	0.003	0.005	0.003
	Std. Dev. (%)	0.0333	0.0384	0.0433	0.0334	0.0294	0.0258	0.0239	0.0295	0.0260
	N	206	189	167	148	131	116	97	76	41
All	M.R.P. (%)	0.004	0.007	0.008	0.008	- .002	- .002	0.004	0.005	0.002
	Std. Dev. (%)	0.0362	0.0382	0.0416	0.0338	0.0287	0.0251	0.0233	0.0292	0.0252
	N	264	228	196	165	144	129	104	79	43
<u>POTASSIUM</u>										
Flame Photometry	M.R.P. (%)	- .003	- .009	- .001	0.002	0.010	0.017	0.026	0.051	0.000
	Std. Dev. (%)	0.0681	0.0741	0.0632	0.0595	0.0812	0.0761	0.0859	0.1158	0.0335
	N	207	188	167	147	132	116	97	72	41

Table C1. Temporal Effect of Interstate Laboratory Participation in
CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>POTASSIUM (continued)</u>										
All	M.R.P. (%)	-0.006	-0.006	0.000	0.000	0.005	0.012	0.026	0.052	-0.002
	Std. Dev. (%)	0.0811	0.0778	0.0637	0.0583	0.0804	0.0767	0.0849	0.1153	0.0327
	N	271	225	195	164	146	127	103	76	44
<u>CHLORIDE</u>										
Scales	M.R.P. (%)	0.006	0.000	0.001	0.013	0.007	0.008	0.024	-----	-----
	Std. Dev. (%)	0.0607	0.0482	0.0438	0.0608	0.0523	0.0602	0.0448	-----	-----
	N	138	110	94	76	62	52	43	-----	-----
All	M.R.P. (%)	0.004	0.001	0.001	0.005	0.000	0.002	0.019	0.009	-----
	Std. Dev. (%)	0.0665	0.0593	0.0434	0.0562	0.0518	0.0555	0.0407	0.0383	-----
	N	381	264	205	153	123	102	76	49	-----
<u>CALCIUM</u>										
SMA	M.R.P. (%)	0.001	0.014	0.002	-0.005	-0.005	-0.011	-0.014	-0.031	-----
	Std. Dev. (%)	0.0584	0.0775	0.0402	0.0636	0.0425	0.0434	0.0551	0.0679	-----
	N	126	99	84	76	65	52	48	40	-----
All	M.R.P. (%)	0.010	0.012	0.000	-0.004	-0.004	-0.009	-0.003	-0.031	-0.003
	Std. Dev. (%)	0.1070	0.0945	0.0560	0.0662	0.0550	0.0570	0.0557	0.0676	0.0376
	N	406	277	220	180	143	109	84	50	30

Table C1. Temporal Effect of Interstate Laboratory Participation in
CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>HEMATOCRIT</u>										
Micro-Hematocrit	M.R.P. (%)	0.002	0.005	0.006	0.005	0.010	0.004	0.001	0.004	-----
	Std. Dev. (%)	0.0731	0.0543	0.0455	0.0512	0.0378	0.0398	0.0699	0.0972	-----
	N	188	166	152	129	111	92	69	48	-----
All	M.R.P. (%)	0.002	0.001	0.006	0.005	0.010	0.004	0.002	0.004	-----
	Std. Dev. (%)	0.0951	0.0756	0.0452	0.0508	0.0378	0.0398	0.0699	0.0972	-----
	N	263	214	194	159	128	101	69	48	-----
<u>RED BLOOD CELLS</u>										
All	M.R.P. (%)	0.022	0.003	0.026	0.036	0.027	0.009	0.017	0.023	-----
	Std. Dev. (%)	0.1336	0.1042	0.1127	0.0961	0.0974	0.1084	0.1148	0.1187	-----
	N	302	247	198	144	110	90	52	39	-----
<u>HEMOGLOBIN</u>										
Cyanmethamoglobin	M.R.P. (%)	-.023	-.028	-.013	-.018	0.000	-.012	-----	-----	-----
	Std. Dev. (%)	0.0786	0.0957	0.0559	0.0494	0.0577	0.0359	-----	-----	-----
	N	158	123	89	69	53	39	-----	-----	-----
Coulter S.	M.R.P. (%)	0.036	-.002	0.010	0.003	0.013	-----	-----	-----	-----
	Std. Dev. (%)	0.2182	0.0300	0.0344	0.0216	0.0240	-----	-----	-----	-----
	N	79	58	46	36	29	-----	-----	-----	-----

Table C1. Temporal Effect of Interstate Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)											
	0	3	6	9	12	15	18	21	24			
<u>HEMOGLOBIN (continued)</u>												
All	M.R.P. (%)	0.014	0.001	-.001	-.009	0.005	0.005	0.008	-----	-----	-----	-----
	Std. Dev. (%)	0.1750	0.1354	0.0672	0.0467	0.0473	0.1352	0.0945	-----	-----	-----	-----
	N	387	276	203	151	110	74	48	-----	-----	-----	-----

Table C2. Temporal Effect of Volunteer Laboratory Participation in
CDC Proficiency Testing Program

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)																				
	0	3	6	9	12	15	18	21	24												
<u>GLUCOSE</u>																					
AutoAnalyzer	M.R.P. (%)	0.001	0.000	0.007	-0.006	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	
	Std. Dev. (%)	0.0752	0.0595	0.0526	0.0512	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651	0.0651
	N	70	63	54	50	43	35	35	35	35	35	35	35	35	35	35	35	35	35	35	35
All	M.R.P. (%)	0.055	0.030	0.031	0.026	0.034	0.034	0.034	0.034	0.034	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021	
	Std. Dev. (%)	0.2827	0.1969	0.1349	0.1525	0.1576	0.1576	0.1576	0.1576	0.1576	0.0970	0.0970	0.0970	0.0970	0.0970	0.0970	0.0970	0.0970	0.0970	0.0970	0.0830
	N	289	199	144	112	83	68	68	68	68	49	49	49	49	49	49	49	49	49	49	49
<u>CREATININE</u>																					
Alkaline Picrate w/o Lloyds	M.R.P. (%)	-0.005	0.005	0.013	0.018	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061	0.061
	Std. Dev. (%)	0.3068	0.5738	0.1977	0.2063	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783	0.2783
	N	69	53	41	36	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
All	M.R.P. (%)	0.001	0.052	-0.003	-0.006	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	
	Std. Dev. (%)	0.3198	0.4481	0.1655	0.1794	0.2286	0.2286	0.2286	0.2286	0.2286	0.2157	0.2157	0.2157	0.2157	0.2157	0.2157	0.2157	0.2157	0.2157	0.2157	0.1363
	N	159	105	83	73	60	46	46	46	46	33	33	33	33	33	33	33	33	33	33	33
<u>UREA NITROGEN</u>																					
AutoAnalyzer	M.R.P. (%)	0.006	0.000	-0.024	0.083	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
	Std. Dev. (%)	0.1472	0.0880	0.0754	0.2614	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517	0.1517
	N	58	49	44	40	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33

Table C2. Temporal Effect of Volunteer Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)											
	0	3	6	9	12	15	18	21	24			
<u>UREA NITROGEN (continued)</u>												
All	0.173	0.098	0.094	0.057	- .014	0.006	-.009	----	----			
M.R.P. (%)	0.4761	0.3548	0.3742	0.4552	0.1956	0.2366	0.1263	----	----			
Std. Dev. (%)	233	153	122	94	71	54	33	----	----			
N												
<u>URIC ACID</u>												
All	0.052	0.005	-.039	-.023	0.038	0.008	-.010	----	----			
M.R.P. (%)	0.7308	0.2059	0.1303	0.1913	0.3154	0.1173	0.0980	----	----			
Std. Dev. (%)	196	138	109	88	67	48	39	----	----			
N												
<u>CHOLESTEROL</u>												
All	0.031	0.032	0.037	0.016	0.038	0.017	0.052	----	----			
M.R.P. (%)	0.1403	0.1386	0.1653	0.1037	0.1347	0.1292	0.0951	----	----			
Std. Dev. (%)	221	142	104	73	58	42	35	----	----			
N												
<u>TOTAL PROTEIN</u>												
All	0.008	0.011	0.011	0.008	-.005	0.001	0.005	----	----			
M.R.P. (%)	0.0794	0.0601	0.0673	0.0582	0.0709	0.0633	0.0418	----	----			
Std. Dev. (%)	200	136	100	77	60	44	33	----	----			
N												

Table C2. Temporal Effect of Volunteer Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)										
	0	3	6	9	12	15	18	21	24		
<u>SODIUM</u>											
Flame Photometry	M.R.P. (%)	0.000	0.002	0.000	-0.006	0.000	0.000	-0.002	0.000	0.006	-----
	Std. Dev. (%)	0.0537	0.0371	0.0275	0.0262	0.0302	0.0276	0.0243	0.0243	0.0295	-----
	N	105	86	75	58	50	46	39	39	29	-----
All	M.R.P. (%)	-0.002	0.002	0.000	-0.005	0.000	0.000	0.000	0.000	0.005	-----
	Std. Dev. (%)	0.0501	0.0396	0.0269	0.0244	0.0281	0.0263	0.0227	0.0227	0.0287	-----
	N	144	111	92	70	61	56	45	45	32	-----
<u>POTASSIUM</u>											
Flame Photometry	M.R.P. (%)	-0.032	-0.003	-0.013	0.004	-0.004	-0.009	0.011	0.011	-0.006	-----
	Std. Dev. (%)	0.1276	0.0615	0.0603	0.0607	0.0440	0.0772	0.0631	0.0631	0.0797	-----
	N	103	86	74	59	50	46	40	40	29	-----
All	M.R.P. (%)	-0.024	-0.005	-0.010	0.004	-0.003	-0.008	0.007	0.007	-0.006	-----
	Std. Dev. (%)	0.1204	0.0596	0.0693	0.0564	0.0414	0.0706	0.0593	0.0593	0.0761	-----
	N	146	110	91	70	61	56	46	46	32	-----
<u>CHLORIDE</u>											
All	M.R.P. (%)	0.007	0.000	-0.004	-0.003	0.000	0.002	0.000	0.000	-----	-----
	Std. Dev. (%)	0.2288	0.0564	0.0427	0.0532	0.0380	0.0300	0.0376	0.0376	-----	-----
	N	170	113	89	67	53	48	38	38	-----	-----

Table C2. Temporal Effect of Volunteer Laboratory Participation in CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>CALCIUM</u>										
All	M.R.P. (%)	- .006	- .008	- .001	0.002	0.000	- .035	-----	-----	-----
	Std. Dev. (%)	0.1587	0.1162	0.1340	0.0915	0.0603	0.1010	-----	-----	-----
	N	184	121	84	68	54	37	-----	-----	-----
<u>HEMATOCRIT</u>										
46	Micro-Hematocrit	M.R.P. (%)	0.000	0.007	0.007	- .001	0.004	0.008	0.007	0.006
		Std. Dev. (%)	0.0522	0.0682	0.0796	0.0753	0.0544	0.0722	0.0373	0.0415
		N	174	148	122	108	98	85	67	51
	All	M.R.P. (%)	- .004	0.004	0.003	- .001	0.004	0.008	0.007	0.006
		Std. Dev. (%)	0.0773	0.0682	0.1194	0.0753	0.0544	0.0722	0.0373	0.0415
		N	186	151	132	108	98	85	67	51
<u>RED BLOOD CELLS</u>										
	All	M.R.P. (%)	0.038	0.000	0.009	0.059	0.029	0.050	-----	-----
		Std. Dev. (%)	0.2372	0.1907	0.1594	0.1766	0.1533	0.1672	-----	-----
		N	171	120	90	72	58	43	-----	-----
<u>HEMOGLOBIN</u>										
	Hemoglobin	M.R.P. (%)	- .005	- .004	0.001	- .001	- .005	- .010	-0.001	-----
	Cyanmethemoglobin	Std. Dev. (%)	0.0718	0.0683	0.0768	0.0753	0.0431	0.0591	0.0850	-----
		N	150	125	102	80	71	57	44	-----

Table C2. Temporal Effect of Volunteer Laboratory Participation in
 CDC Proficiency Testing Program (continued)

Test Type & Method	Elapsed Time in Proficiency Testing Program Since First Shipment (Months)									
	0	3	6	9	12	15	18	21	24	
<u>HEMOGLOBIN (continued)</u>										
All										
M.R.P. (%)	0.043	0.017	0.007	--.001	--.016	--.001	0.009	--.002	-----	-----
Std. Dev. (%)	0.2112	0.1761	0.1451	0.0939	0.0472	0.0792	0.0901	0.0410	-----	-----
N	244	168	133	100	84	66	49	32	-----	-----

Table C3 lists the most frequently used analytical methods for each of the constituents in the CDC proficiency testing data base.

Table C3. Most Popular Analytical Methods

<u>Constituent</u>	<u>Methods</u>
Glucose	(1) AutoAnalyzer; (2) SMA
Creatinine	(1) Alkaline picrate (w/o Lloyds); (2) AutoAnalyzer
Urea Nitrogen	(1) AutoAnalyzer; (2) SMA
Uric Acid	(1) AutoAnalyzer; (2) SMA; (3) Phosphotungstate-carbonate
Total Protein	(1) AutoAnalyzer; (2) SMA; (3) Biuret, serum standard; (4) Refractive index
Cholesterol	(1) AutoAnalyzer; (2) SMA; (3) Direct serum (acetic anhydride)
Triglyceride	(1) Acetylacetone reagent; (2) Enzymatic
Calcium	(1) AutoAnalyzer; (2) SMA; (3) Atomic Absorption
Sodium	(1) Flame Photometry; (2) SMA
Potassium	(1) Flame Photometry; (2) AutoAnalyzer; (3) SMA
Chloride	(1) Cotlove Titrator; (2) AutoAnalyzer; (3) SMA
Hematocrit	(1) Micro-Hematocrit
Red Blood Cell Count	(1) Coulter counter
Hemoglobin	(1) Coulter counter, model S; (2) Cyanmethemoglobin (manual)

Appendix D. Measuring Laboratory Performance in the CDC Proficiency Testing Program

In very general terms, clinical laboratory analyses fall into two procedural categories; quantitative tests (chemistry, hematology, differential blood films, some seriology and the radioisotopes) and qualitative tests (bacteriology, cytology, urinalysis, some serology, and most of immunohematology, and toxicology). The next few paragraphs discuss the CDC proficiency testing program in these two basic categories.

Quantitative Tests

Reference laboratories, usually 100 or more, are used as prescribed by regulation (Federal Register No. 253, Vol. 33, December 31, 1968, p. 20047). Criteria for satisfactory performance for each sample are enumerated in paragraph 74.42; they basically require the use of three subsets of limits. The first subset encompasses the central 95 percentile calculated from the results of all applicant and licensed laboratories. Laboratory results which are "obviously deviant from the other results" are not used. No specific criteria or procedures are provided for determining which values are "obviously deviant," however, there is a stipulation that, for any sample, no more than five percent of the results are to be discarded on the basis of being "obviously deviant." The second subset of limits is determined from the reference laboratory results; they encompass all the results and simply consist of the highest and lowest values (the same procedure is followed for outliers as was just described). The final subset of limits (for clinical requirements) is centered on the median reference laboratory result. "They encompass one-half of the normal population range, when the sample is in the normal range, or one-half of the appropriate range in the case of a sample which is outside the normal range." The normal population range or other appropriate range is determined by the Secretary, DHEW. The results for applicant and licensee laboratories must fall between the lowest lower limit and highest upper limit of the three subsets of limits just described.

For chemistry and hematology, a score of three is assigned to all laboratory results within the tightest combination of limits around the target value. A score of two is assigned to all laboratory results within the next tightest combination of limits (but not within the tightest combination), and a score of one is assigned to all laboratory results falling within the outermost limits determined from the three subsets of limits (but not within the two previously described limits). All results falling outside of the outermost limits are assigned a score of -1. All laboratory results having a score of one or greater are considered acceptable.

The scoring methods for differential blood films and non-syphilis serology are slightly more complicated. They can be found in a report entitled "Development of Grading Methods," a report of the Study Group

for the Development of Grading Methods, presented at the First National Conference on Proficiency Testing, Atlanta, Georgia, October 4-6, 1971.

Qualitative Tests

Qualitative tests differ from the quantitative tests in that in addition to using a select group of three or more reference laboratories, the samples are also sent to ten or more referee laboratories. The reference laboratories analyze the samples in great detail while the referee laboratories handle the samples in the same manner as do the licensed laboratories being tested.

The grading procedures for the qualitative tests are much more complex than those which were described for the quantitative tests. The reader is directed to pages 20047-8, in the December 31, 1968 Federal Register for a complete discussion of this subject.

Appendix E. Mean Measures of Relative Performance

The tables included in this appendix list mean measures of relative performance (MRP) which have been found to be significantly different from each other or from other subgroups. A separate table appears for each of the 14 constituents. In each instance, the difference between individual means (calculated from all samples covered in the two-year period) or subgroups was significant at the 0.001 level of probability. This significance was based on the approximate F-test made necessary by significantly heterogeneous variances ($p < .001$).

Table E1. Glucose

At least 15 laboratories used each of the following six analytical methods. Means which are underscored by the same line are not significantly different from each other. Hence, methods numbered 7, 9, 4, 8 and 6 are in the same subgroup, as are methods 6 and 3; while method 3 is significantly different from methods 7, 9, 4 and 8.

Analytical Method	Mean Relative Performance	Number of Laboratories
#7 - AutoAnalyzer	-1.5%	86
#9 - Other Automated	-1.3%	20
#4 - Direct Serum	- .2%	47
#8 - SMA or Neocuproine	0.2%	93
#6 - Other Manual	1.7%	21
#3 - Ortho Toluidine	5.5%	17
	<u>7, 9, 4, 8, 6, 3</u>	

Table E2. Creatinine

There were five methods which were used by more than 15 laboratories each. These five methods fall into four subgroups with only one instance of commonality.

Analytical Method	Mean Relative Performance	Number of Laboratories
#6 - Other Manual	-22.3%	17
#2 - Alkaline Picrate, w/o Lloyds	- 5.1%	149
#8 - SMA	- 3.5%	22
#7 - AutoAnalyzer	- 3.3%	53
#1 - Alkaline Picrate, Lloyds	16.1%	29
	<u>6 2 8, 7 1</u>	

Table E3. Urea Nitrogen

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Laboratories</u>
#8 - SMA	-1.9%	97
#1 - Urease-Nesolerization	- .7%	15
#7 - AutoAnalyzer	- .6%	80
#2 - Urease-Berthelot	2.8%	23
#3 - Diacetyl Reaction	3.6%	35
#6 - Other Manual	7.2%	15
#9 - Other Automated	39.0%	16
#5 - Kits	39.7%	17

8, 1, 7 2, 3 6 9, 5

Table E4. Uric Acid

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#6 - Other Manual	-8.6%	19
#1 - Phosphotungstate-cyanide	-3.8%	18
#2 - Phosphotungstate-carbonate	-3.0%	73
#8 - SMA	-1.6%	94
#7 - AutoAnalyzer	0.7%	64

6 1, 2 8 7

Table E5. Calcium

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#7 - AutoAnalyzer	-1.0%	32
#5 - Flame Photometer	- .9%	24
#8 - SMA	- .1%	112
#2 - EDTA Titrimetric	0.4%	48
#3 - EDTA Colorimetric	1.2%	27
#4 - Atomic Absorption	2.0%	27
#6 - Other Manual	2.4%	48
#9 - Other Automated	2.6%	20

7, 5 8 2 3 4, 6, 9

Table E6. Chloride

Analytical Method	Mean Relative Performance	Number of Observations
#8 - SMA	-.5%	20
#2 - Cotleve Titrator	-.4%	54
#3 - Other Titrimetric	-.1%	44
#1 - Schales	0.6%	111
#7 - AutoAnalyzer	0.7%	33
#5 - Kits	1.0%	26
	<u>8, 2</u> <u>3</u> <u>1, 7</u> <u>5</u>	

Table E7. Total Protein

Analytical Method	Mean Relative Performance	Number of Observations
#7 - AutoAnalyzer	-.7%	32
#2 - Biuret, Serum	-.6%	79
#8 - SMA	-.5%	114
#5 - Refractive Index	0.6%	29
	<u>7, 2, 8</u> <u>5</u>	

Table E8. Potassium

Analytical Method	Mean Relative Performance	Number of Observations
#9 - Other Automated	-4.8%	16
#7 - AutoAnalyzer	-.3%	16
#1 - Flame Photometry	0.4%	169
#8 - SMA	0.9%	17
	<u>9</u> <u>7, 1, 8</u>	

Table E9. Sodium

Analytical Method	Mean Relative Performance	Number of Observations
#9 - Other Automated	-.7%	15
#1 - Flame Photometry	0.3%	168
#8 - SMA	0.5%	18
#7 - AutoAnalyzer	0.6%	18
	<u>9</u> <u>1</u> <u>8, 7</u>	

Table E10. Cholesterol

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#7 - AutoAnalyzer Extract	-3.3%	67
#3 - Ferric-Sulfuric	-2.5%	18
#8 - SMA	-2.5%	97
#5 - Direct Serum, Ferric-Sulfuric	-2.4%	21
#4 - Direct Serum, Acetic Anhydride	-2.1%	87
#6 - Other Manual	- .1%	48
#9 - Other Automated	0.6%	18
	<u>7, 3, 8, 5</u> 4 <u>6, 9</u>	

Table E11. Triglycerides

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#1 - Acetylocetone Reagent	1.2%	132
#2 - Chromotropic Acid	10.0%	18
#5 - Other	13.7%	28
	<u>1</u> <u>2</u> <u>5</u>	

Table E12. Hemoglobin

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#9 - Other Automated	-2.1%	47
#1 - Cyanmethemoglobin	-1.9%	124
#6 - Coulter S	2.1%	75
#8 - SMA	2.3%	20
#5 - Other Manual	21.1%	16
	<u>9, 1</u> <u>6, 8</u> <u>5</u>	

Table E13. Red Blood Cell Count

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#3 - SMA	-0.6%	16
#2 - Coulter	0.7%	126
#5 - Other Automated	2.6%	15
#1 - Manual	4.9%	30
	<u>3</u> <u>2</u> <u>5</u> <u>1</u>	

Table E14. Hematocrit

<u>Analytical Method</u>	<u>Mean Relative Performance</u>	<u>Number of Observations</u>
#3 - Other	-4.8%	15
#2 - Micro Hematocrit	0.4%	159
	<u>3</u> <u>2</u>	

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBSIR 75-197	2. Gov't Accession No.	3. Recipient's Accession No.
4. TITLE AND SUBTITLE Clinical Laboratory Performance Analysis Using Proficiency Test Statistics		5. Publication Date	6. Performing Organization Code
7. AUTHOR(S) Peter W. Finkel and John W. Rowen		8. Performing Organization NBSIR 73-197	
9. PERFORMING ORGANIZATION NAME AND ADDRESS NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234		10. Project/Task/Work Unit No. 4314457	11. Contract/Grant No.
12. Sponsoring Organization Name and Address Department of Health, Education and Welfare Office of the Assistant Secretary for Planning and Evaluation Division of Health Evaluation, Washington, D.C. 20201		13. Type of Report & Period Covered Final 1/71 - 3/73	14. Sponsoring Agency Code
15. SUPPLEMENTARY NOTES This report supersedes COM-73-11253			
16. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.) <p>The proficiency testing aspects of the Clinical Laboratory Improvement Act of 1967 program were assessed. The overall ability of licensed or volunteer laboratories to accurately determine mean values for any of the 13 constituents was not significantly altered during the first two years of program operation. However, the variability of the laboratories has decreased over the two-year period. It appears that the program has increased the consistency of laboratory performance.</p> <p>The general level of laboratory capability seemed to be independent of involvement in state-supported or voluntary proficiency testing programs other than the CDC program and of whether the working supervisor had a B.S., M.S., Ph.D., or M.D. degree. Choice of analytical method did significantly affect performance. Although insufficient evidence was available to make a definitive statement, the data do not appear to support arguments favoring establishment of method-dependent reference group target values.</p> <p>Finally, it appears that consideration should be given to alternative sampling methods, such as reduced or skip-lot sampling, for those constituents which appear to present no analytical challenge to the licensed and volunteer laboratories. Greater emphasis might then be placed on those constituents which give the laboratories the most difficulty (cholesterol and creatinine, for example).</p>			
17. KEY WORDS (Alphabetical order, separated by semicolons) Accuracy; CLIA '67; clinical chemistry; hematology; microbiology; precision; proficiency testing			
18. AVAILABILITY STATEMENT <input checked="" type="checkbox"/> UNLIMITED. <input type="checkbox"/> FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NTIS.		19. SECURITY CLASS (THIS REPORT) UNCLASSIFIED	21. NO. OF PAGES
		20. SECURITY CLASS (THIS PAGE) UNCLASSIFIED	22. Price

