International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Inter-Laboratory Comparability of Clinical Chemistry Testing: A New Perspective

Perez Quartey[a]*, Ruth Anokye[b], Michael Frimpong[c]

[a,b,c]*Department of Medical Laboratory Technology, Garden City University College, Kenyase, Ghana*

[a]*Email: quartey.perez@gcuc.edu.gh*

[b]*Email: obaaabrafi29@gmail.com*

[c]*Email: mfrimpongboadu@gmail.com*

Abstract

Background/Aim: Existing external quality assurance proficiency testing may not reflect routine 'real-world' daily testing proficiency. The overall aim of the study was to conduct an inter-laboratory comparability testing of some routinely measured clinical chemistry analytes by selected private medical laboratories using the mentor-adept approach. **Methodology:** Aliquots of freshly separated plasma from a single individual (1ml) each were sent to each of the selected labs and a Teaching hospital for same-day testing. All samples were sent as ''blinded-samples'' (labelled with anonymous names, ages and attached lab request forms with clinical diagnoses) so that they would be tested as real patient samples. **Results:** All the labs met the acceptability criteria range for both the *z-score* and Precision Index. However, many of the labs had their total analytical errors for the tests outside the allowable total error ranges with both European and CLIA recommendations. **Conclusion:** There are relative similarities in the *z-score* (inter-laboratory bias) and Precision Index (inter-laboratory precision) among the labs. However, many of the labs did not meet recommended analytical goals for total analytical errors on individual samples run in a day. In the light of the findings, it is highly recommended that though laboratories should be aware of conduction of periodic external quality assurance exercises, such exercises should be done using "blinded-samples" as utilized in this current study without prior notification of the day and time of testing.

Keywords: Inter-laboratory; Mentor-adept; External Quality Assurance; z-score; Precision Index; Total Analytical Error.

---

* Corresponding author.

1. Introduction

Quality of clinical chemistry testing is very critical in diagnostic and overall patient assessment by expressing reliability in testing results. Clinical chemistry testing presents as a very critical aspect of general laboratory testing and it takes highly automated measurement methods and systems, combined with advanced information technologies to accomplish this all-important task [1]. The overall uncertainty of the high-volume measurement methods in clinical chemistry has decreased substantially in recent decade tough bias still remains as a significant challenge. A key component of laboratory quality management system is Quality Assurance (QA) which includes both Internal Quality control (done on regular daily basis) and External Quality Assurance (EQA). EQA is done periodically, in collaboration with a higher laboratory authority through Proficiency Testing (PT) or Inter-lab comparability (ILC) testing. with PT being the ideal method of choice in evaluating laboratory performance [2, 3, 4, 5]. In cases where for some reasons there is lack of proficiency testing schemes, inter-laboratory comparisons (ILCs) are preferred by accreditation bodies [5]. Typical method of assessment of ILC results have been described with standard statistical protocols [4, 6, 7] with the use of 'reference' materials prepared by the QA provider for measurements. When it is not possible to apply this, the mentor-adept approach of inter-laboratory comparison may be used [1]. Measurement results of analytes from the expert (mentor-lab), ''considered'' to be devoid of bias [1] are compared with that of the participating laboratories. The inter-laboratory analysis of the results is done using recommended statistical methods. Statistically stringent or robust analysis ''consensus'' values calculated from participant results are recommended for the estimation of an assigned value and are typically used for calculating the *z-score*, precision index (PI) and Total analytical error (TE) [4, 5, 6, 8]. Presently, we are not aware of any external quality assurance exercise conducted for clinical chemistry testing by medical laboratories in Ghana.

1.1. Aim of the Study

The overall aim of the study was to conduct an inter-laboratory comparability testing of some routinely measured clinical chemistry analytes by some selected private medical labs in Kumasi (the second largest city in the country) using the mentor-adept approach of performance evaluation.

2. Materials and Methods

2.1. Ethical consideration

Approval for the research was received from the Kwame Nkrumah University of Science and Technology Research Committee.

2.2. Study design

Using the Mentor-Adept approach, the cross-sectional study was conducted on six (6) selected private laboratories.

2.3. Mentor lab

Komfo Anokye Teaching Hospital Chemical Pathology laboratory.

2.4. Adept-labs

Six (6) selected private labs in the Kumasi metropolis. The locations of two of the labs are relatively farther away from the mentor-lab whiles the other four are in the vicinity of the mentor-lab. All labs were assigned alphabetical codes for ethical anonymity.

2.5. Study protocol and procedures

After thoroughly explaining the aim, objectives and procedures of the study, venous blood was collected from a single healthy donor. Freshly separated plasma was aliquoted into clean eppendorf tubes. Seven (7) of the 1ml plasma aliquots were sent to each of the selected labs for same-day testing and reporting. All samples were sent as ‘‘blinded-samples’’ (labeled with anonymous names, ages and attached lab request forms with clinical diagnoses) so that they would be tested as real patient samples.

2.6. Measurements

Sodium, Potassium, Urea, Creatinine, Total Cholesterol, HDL, Triglycerides.

2.7. Data analysis

Comparability of precision, accuracy and total analytical error were assessed using standard lab quality assurance statistical methods. Inter-laboratory comparability of analytical bias was assessed by calculation of ‘‘Consensus-based’’ Standard Deviation Index (z-score), using robust calculation involving exclusion of outlier results [4].

$$z - score = \frac{\text{x(lab mean)} - \text{X(consensus mean)}}{\text{consensus standard deviation}}$$

Classification criteria:

0 to ±1 = Satisfactory

>1 to 2 = Acceptable

± > 2 = Unacceptable

Inter-laboratory comparability of imprecision was assessed calculation of ‘‘Consensus-based’’ Precision Index [4].

$$PI = \frac{\text{lab standard deviation}}{\text{consensus standard deviation}}$$

Classification criteria:

0 to ±1 = Satisfactory

>1 to 2 = Acceptable

± > 2 = Unacceptable

Inter-laboratory comparability of overall analytical performance was assessed with total analytical error and comparing it with two recommended specifications of allowable total error (TEa) [4, 9, 10].

The following methods were used in evaluation of performance:

a. Total analytical error calculated as 1.65 x CV of lab + % bias, where % bias is;

$$\frac{\text{lab mean} - \text{mentor lab mean}}{\text{mentor lab mean}} \text{ x } 100$$

This error is compared with European recommended allowable total error for each analyte by Ricos and his colleagues 2014 [9].

b. Lab means compared with CLIA [10] recommended allowable total error calculated as mentor mean (target value) ± recommended % bias.

3. Results

3.1. General characteristics, Z-score and PI scores of the labs

The general characteristics, *Z*-scores and *PI*-scores of the participant labs are shown on results of the *z-score* distributions in the participating labs are illustrated on table 1, 2 and 3 respectively. The results showed that all the mean results of tests were within the acceptability criteria range (± 2 standard deviation points from the consensus mean). All the mean results of Lab A were in the satisfactory range. The mean results that moved from the satisfactory range into the acceptable range were urea and creatinine for Lab B, sodium and creatinine for Lab C, HDL for Labs D and F, and both urea and triglyceride for Lab E. Additionally, all the participant labs met the acceptability criteria for precision index (± 2). A detailed analysis however showed that the PI moved from the satisfactory range to the acceptable range for sodium (Lab B) potassium (Lab C) and both LDL and HDL for Lab D.

Table 1: General Characteristics of the study laboratories.

Lab	Sodium	Potassium	Urea	Creatinine	Cholesterol	HDL	Triglycerides
Mentor-lab	149.2 ± 0.36	4.2 ± 0.11	2.25 ± 0.02	71.7 ± 1.53	3.22 ± 0.12	0.92 ± 0.01	1.14 ± 0.01
A	146.7 ± 2.18	4.1 ± 0.08	3.8 ± 0.42	75.5 ± 7.13	3.77 ± 0.27	1.04 ± 0.16	1.61 ± 0.28
B	121.5 ± 21.85	3.9 ± 0.31	2.5 ± 0.01	137.6 ± 1.00	3.21 ± 0.01	0.95 ± 0.01	1.34 ± 0.03
C	105.1 ±1.76	3.5 ± 0.43	3.8 ± 0.48	71.4 ± 1.55	3.91 ± 0.11	1.13 ± 0.01	1.78 ± 0.06
D	145.5 ± 0.86	4.1 ± 0.02	3.6 ± 0.71	87.9 ± 10.31	4.87 ± 0.43	1.83 ± 0.46	1.73 ± 0.10
E	137.9 ± 1.35	3.8 ± 0.10	5.3 ± 0.05	80.8 ± 16.19	3.79 ± 0.41	1.37 ± 0.30	0.96 ± 0.05
F	143.3 ± 1.60	4.3 ± 0.30	4.7 ± 0.17	60.9 ± 3.12	3.80 ± 0.30	0.69 ± 0.07	1.27 ± 0.08
Consensus Mean	133.3 ± 17.55	4.0 ± 0.34	4.0 ± 0.96	85.7 ± 26.18	3.88 ± 0.57	1.17 ± 0.43	1.49 ± 0.10

Results are expressed as mean ± SD in mmol/l.

Table 2: Z-scores of the study labs

Lab	Sodium	Potassium	Urea	Creatinine	Cholesterol	HDL	Triglycerides
A	0.8	0.3	-0.2	0.4	-0.2	-0.3	0.4
B	-0.7	-0.3	-1.6$^{\alpha}$	2.0$^{\alpha}$	-1.2	-0.5	-0.5
C	-1.6$^{\alpha}$	-1.5$^{\alpha}$	-0.2	-0.6	0.1	-0.1	0.9
D	0.7	0.3	-0.4	0.1	1.7	1.5$^{\alpha}$	0.7
E	0.3	0.6	1.4$^{\alpha}$	0.2	-0.2	0.5	-1.6$^{\alpha}$
F	0.6	0.9	0.7	1.0	-0.1	-1.1$^{\alpha}$	-0.7

Scores with no superscript means values are in satisfactory range (less than ± 1.0); lab means with superscript ($^{\alpha}$) means values

are in acceptable range (± 1.1 – 2.0).

Table 3: Precision Indexes (PI) of the study labs

LAB	Sodium	Potassium	Urea	Creatinine	Cholesterol	HDL	Triglycerides
A	0.1	0.2	0.02	0.3	0.5	0.4	0.8
B	1.3 $^{\alpha}$	0.9	0.4	0.04	0.2	0.02	0.1
C	0.1	1.3 $^{\alpha}$	0.5	0.1	0.02	0.02	0.2
D	0.05	0.1	0.7	0.4	0.8	1.1 $^{\alpha}$	0.3
E	0.1	0.3	0.1	0.6	0.7	0.7	0.2
F	0.1	0.9	0.2	0.1	0.5	0.2	0.2

Scores with no superscript means PI values are in satisfactory range (less than ± 1.0); Scores with superscript ($^{\alpha}$) means PI values are in acceptable range (± 1.1 – 2.0).

Table 4: Total analytical errors of the study labs compared by the European recommendations.

Analyte	Lab A	Lab B	Lab C	Lab D	Lab E	Lab F	TEa (%)
Sodium							
CV (%)	1.5	18.0	1.7	0.6	1.0	1.1	
%Bias	1.7	18.6	29.6	2.5	7.6	4.0	
TE (%)	4.2 β	48.3 β	32.4 β	3.5 β	9.3 β	5.8 β	1.0
Potassium							
CV (%)	2.0	7.9	12.3	0.5	2.6	6.9	
%Bias	2.4	7.1	16.7	2.4	9.5	2.4	
TE (%)	5.7 β	20.1 β	40.0 β	3.2	13.8 β	13.8 β	5.61
Urea							
CV (%)	11.1	0.4	12.6	19.7	0.9	3.6	
%Bias	68.9	11.1	68.9	56.5	130.4	104.3	
TE (%)	87.2 β	11.7	96.3 β	89.0 β	131.5 β	110.2 β	15.55
Creatinine							
CV (%)	9.4	0.7	2.2	11.7	20.0	5.1	
%Bias	5.3	92.0	0.4	22.6	12.7	15.1	
TE (%)	20.8 β	93.2 β	4.0	41.9 β	45.7 β	23.5 β	8.87
Cholesterol							
CV (%)	7.2	0.3	2.8	8.8	10.8	7.9	
%Bias	17.1	0.3	21.4	51.2	17.7	18.0	
TE (%)	29.0 β	0.8	26.0 β	65.7 β	35.5 β	31.0 β	9.01
HDL							
CV (%)	15.4	1.1	0.9	25.1	21.9	10.1	
%Bias	3.3	3.2	22.8	98.9	48.9	25.0	
TE (%)	28.7 β	5.0	24.3 β	140.3 β	85.0 β	41.6 β	11.63
Triglycerides							
CV (%)	17.4	2.2	3.4	5.8	5.2	6.3	
%Bias	41.2	17.5	56.1	51.7	18.8	11.4	
TE (%)	69.9 β	21.1	61.7 β	61.3 β	27.4 β	21.8	25.99

TEa: Allowable total error; β: Total analytical error beyond the allowable total error range for the analyte.

Table 5: Total analytical errors (TE) of the study labs compared by CLIA 2019 recommendations.

Analyte (mmol/l)	Lab A	Lab B	Lab C	Lab D	Lab E	Lab F	Mentor- lab mean	TEa limits
Sodium	146.7	121.5 $^{\beta}$	105.1 $^{\beta}$	145.5	137.9 $^{\beta}$	143.3 $^{\beta}$	149.2	145 - 153
Potassium	4.1	3.9	3.5 $^{\beta}$	4.1	3.8 $^{\beta}$	4.3	4.2	3.9 - 4.5
Urea	3.8 $^{\beta}$	2.5	3.8 $^{\beta}$	3.6 $^{\beta}$	5.3 $^{\beta}$	4.7 $^{\beta}$	2.3	2.09 - 2.51
Creatinine	75.5	137.6 $^{\beta}$	71.4	87.9 $^{\beta}$	80.8 $^{\beta}$	60.9 $^{\beta}$	70.0	63.0 – 77.0
Cholesterol	3.8 $^{\beta}$	3.2	3.9 $^{\beta}$	4.9 $^{\beta}$	3.8 $^{\beta}$	3.8 $^{\beta}$	3.21	2.89 – 3.53
HDL	1.04	0.95	1.13 $^{\beta}$	1.83 $^{\beta}$	1.37 $^{\beta}$	0.69 $^{\beta}$	0.92	0.74 – 1.10
Triglycerides	1.61 $^{\beta}$	1.3	1.8 $^{\beta}$	1.70 $^{\beta}$	0.96 $^{\beta}$	1.27	1.14	0.97 – 1.31

TEa: Allowable total error (CLIA, 2019)**,** $^{\beta:}$ Total analytical error beyond the allowable total error range for the analyte.

Table 6: Comparison of overall performance of the study labs based on European and CLIA total error specifications.

Study Lab	European recommendation	CLIA recommendation
Lab A	0 (0.0%)	4 (57.1%)
Lab B	4 (57.1%)	5 (71.4%)
Lab C	1 (14.3%)	1 (14.3%)
Lab D	1 (14.3%)	2 (28.6%)
Lab E	0 (0.0%)	0 (0.0%)
Lab F	2 (28.6%)	2 (28.6%),

Results are expressed as number (%) of analytes that passed allowable total errors by the recommended specifications.

Overall analytical performance of the study labs

The overall performance was calculated in percentage points of the number of analyte measurements that passed the recommended specifications out of the total number of measured analytes. The overall performance as illustrated in tables 4, 5 and 6 showed that many of the labs had their total analytical errors for the tests outside the allowable total error ranges with only Lab B scoring 4 correct analytical points (57.1%). Labs C, D and F scored 1(14.3%) correct analytical point whilst Labs A and E did not score any point; with the use of the European classification. In comparison to CLIA recommendation however, Lab A scored 4 (57.1%), Lab B scored 5 (71.4%), Lab C scored 1 (14.3%), Labs D and Lab F scored 2 (28.6%), whilst lab E did not score any point.

4. Discussion

Laboratory external quality assessment describes a method or process that allows testing conducted by a laboratory to be compared to that of a source outside the laboratory which may be a peer group of laboratories or a reference laboratory. For accrediting bodies and regulatory agencies, participation in EQA program provides objective evidence of the quality of testing of patients' specimens [9,10]. In general, there is limited data on EQA of clinical chemistry testing in Ghana. Presently, there is no oversight body that organizes External Quality assurance tests of clinical chemistry in Ghana. This is the first study to report on the inter-laboratory comparability of clinical chemistry testing in Ghana, using the mentor-adept approach. Without participation in high-quality inter-laboratory comparability program, a laboratory may be unaware of periodic gradual or sudden changes in performance of the testing system that may be caused by factors such as change in reagents or calibrators, standardization changes, or instrument software changes. Participation in an inter-laboratory comparability program can therefore offer awareness of shifts and trends and help verify the reliability and quality of testing. An inter-laboratory program can also increase the confidence of laboratories in participation in proficiency testing surveys.

4.1. Key Findings of the Current Study

This study typically assessed, the consensus-based *z-score*, precision index, and total analytical error of the selected laboratories in some routinely measured clinical chemistry tests **[**4, 8, 11, 12]. This study presents the likely events of the ‘‘real-life’’ daily results presented to patients and requesting physicians alike when samples are handled as routine patient samples for laboratory analysis and not as samples for an external quality assurance work. In the current study, the results showed that all the mean test results (*z-scores*) and the Precision indexes of the labs were within the acceptability criteria range (± 2). The exclusion of individual laboratory results outliers for calculation of the ‘‘*consensus or peer-group’’* mean and standard deviation implies that interpretation of the *z-score and* PIs typically relies on the statistical assumption that the results are normally distributed and therefore introduces similarities. These parameters do not give an indication of the laboratory's analytical performance but rather an idealized performance of the study group, though they are useful in identifying very questionable analytical deficiencies in the event of a participant lab having a score for an analyte outside the acceptable range. Thus, it has been advocated that in a single round of proficiency testing, providers, participants and end-users should avoid classification and ranking of laboratories based on these scores [13, 14,15]. The comparison of total analytical error for the analytes by the study labs using both European and CLIA recommendations showed that many of the labs had their total analytical errors for the tests outside the allowable total error ranges. A detailed comparison showed that most of the study labs failed with the use of European recommendations as compared to CLIA recommendations. These discrepancies need further investigations for reconciliation and harmonization in the light of differences in specifications and recommendations.

4.2. Key Points of Interest

Anytime a sample is sent to the lab, the requesting physician or patient does not care about the measurement system, testing method or competency of the medical lab technologist doing the testing. All they expect is a reliable result, irrespective of whatever was primarily used to produce the results. For example, a client who walks into a lab with a plasma potassium level of 3.4 mmol/l, irrespective of whatever system, method, time of testing or lab personnel expects a report result within an acceptable range (as close enough to 3.4 mmol/l). External quality assurance typically involves an accredited scheme provider preparing reference materials and sending them to the participating laboratories for measurements [16,17]. The results from the participating labs are then returned to the scheme provider for analysis and feedback reported to the participating labs. The whole process of the proficiency testing is done with the conscious awareness of the participating labs of their involvement in an EQA exercise. External quality assurance programs like proficiency testing and inter-laboratory comparison provide substantial information to measure overall testing performance. However, they do not give a measure of the daily accuracy, reproducibility and overall reliability of individual testing. In our view, EQA surveys may merely provide a snapshot in time and should not be used as a substitute for daily quality. When EQA is conducted with the conscious awareness of the participating labs that the testing is being carried out as part of a formal EQA program, there is a strong likelihood of bias in adhering to strict internal quality control standards with the sole aim of achieving good results. “Acceptable” performance in the EQA testing may therefore not reflect or guarantee testing reliability on individual samples run as real daily work.

Our main issue was to find out the comparability of test results as run by the labs in their real daily testing. Therefore, in the present study no contact was made with the testing personnel to identify any instrumental, methodological or internal quality assurance differences between the labs included in the study. The results of the labs were therefore considered as true results that they would have reported, and dispatched to clients on a ‘‘real-life’’ basis.

4.3. Key Strengths of the Current Study

1. *All samples were sent to the selected labs as "blinded-patient samples". All samples were labelled with anonymous patient names, ages, gender and attached with corresponding request forms with different clinical diagnoses;*

2. *All the selected labs performed testing without a conscious awareness of any inter-laboratory comparability exercise and all results were collected as ‘‘normal patient’’ results as routinely dispatched by the labs. All these show that the results give the most likely test reporting that would have been produced by the labs as part of their routine daily runs.*

5. Conclusion

This study reports the first assessment of inter-laboratory comparability of clinical chemistry testing using the mentor-adept approach in Ghana. The study identified that there are relative similarities in the *z-score* (inter-laboratory bias) and Precision Index (inter-laboratory precision) among the selected labs. However, there are considerable variations in the total analytical errors of testing. The differences in total analytical error as compared to recommended allowable total error criteria may be a better indication of overall analytical performance. In general, many of the laboratories did not meet the recommended analytical goals on individual samples run in a day. These observed analytical deficiencies require critical investigation into the internal quality assurance practices of laboratories in the country.

5.1. Recommendations

It is highly recommended that though laboratories should be aware of conduction of periodic external quality assurance exercises, such exercises should be done using "blinded-samples" without prior notification of the day and time of testing as utilized in this current study.

References

[1]. Theodorsson E Magnusson B, Leito I. Bias in clinical chemistry. *Bioanalysis* 2014; 6(21), 2855–2875.

[2]. ISO, IEC 17025 (2005) General requirements for the competence of testing and calibration laboratories. International Organization for Standardization/International Electrotechnical Commission, Geneva. https://www.iso.org/standard/39883.html. Accessed Feb 2019.

[3]. EA-4/18 (2010) Guidance on the level and frequency of proficiency testing participation. http*://www*.european-accreditation.org/publication/ea-4-18-inf. EA-4/18*:* 2010. Accessed Feb 2019.

[4]. Karkalousos P, Evangelopoulos A (2011). Quality Control in Clinical Laboratories, Applications and Experiences of Quality Control, Prof. Ognyan Ivanov (Ed.), ISBN: 978-953-307-236-4. http://www.intechopen.com/books/applications-and-experiences-of-quality-control/quality-control-inclinical-laboratories. Accessed Feb 2019.

[5]. Szewczak E, Bondarzewski A. Is the assessment of interlaboratory comparison results for a small number of tests and limited number of participants reliable and rational? Accred Qual Assur 2016; 21:91–100.

[6]. ISO 13528 (2005) Statistical methods for use In proficiency testing by interlaboratory comparisons. International Organization for Standardization, Geneva. https://www.iso.org/standard/35664.html. Accessed Feb 2019.

[7]. Hyltoft Peterson P, Ricos C, Stockl D, Libeer JC, Baadenhuijsen H, Fraser C, Thienpont L. ‘‘Proposed guidelines for the internal quality control of analytical results in the medical laboratory.’’ Eur J Clin Chem Clin Biochem 1996; 34:983-999.

[8]. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J, Perich C, Simon M (2014) Desirable specifications for total error, imprecision, and bias, derived from intra- and inter-individual biologic variation. http://www.westgard.com/biodatabase1.htm. Accessed Feb 2019.

[9]. Sciacovelli L, Secchiero S, Zardo L, Plebani M. The role of the External Quality Assessment. Biochemia Medica. 2010; 20(2):160–4.

[10]. Kettelhut MM, Chiodini PL, Edwards H, Moody A. External quality assessment programmes raise standards: evidence from the UKNEQAS parasitology subschemes. J Clin Pathol. 2003; 56(12):927–32.

[11]. ISO 13528 (2015) Statistical methods for use in proficiency testing by interlaboratory comparisons. International Organization for Standardization, Geneva. https://www.iso.org/standard/56125.html. Accessed Feb 2019.

[12]. 2019 CLIA proposed changes to PT acceptable limits. https://www. westgard.com/2019-clia-changes.htm. Accessed April 2019.

[13]. Fearn T, Fisher S, Thompson M, Ellison SLR. “A decision theory approach to fitness-for purpose in analytical measurement”, Analyst 2002; 127:818-824.

[14]. Kristensen GB, Meijer P. Interpretation of EQA results and EQA-based trouble shooting. Biochem Med 2017 27: 49-62.

[15]. Sareen R. Illuminating Z-Score in External Quality Assessment for Medical Laboratory. Health Care Current Review 2018, 6:3.

[16]. ISO/IEC 17043:2010. Conformity assessment – General requirements for proficiency testing. Geneva: International Organization for Standardization; 2010.

[17]. ISO 15189:2013. Medical laboratories – Requirements for quality and competence. International Organization for Standardization; 2013.