sciendo Revista Română de Medicină de Laborator 2025;33(2):119–124 RESEARCH ARTICLE
DOI:10.2478/rrlm-2025-0012

Pilot study on laboratory sample exchange: Evaluation and key findings

Oana Roxana Oprea[1,2], Diana Ionela Handolescu[2*], Minodora Dobreanu[1,2]

1. Department of Laboratory Medicine, George Emil Palade University of Medicine, Pharmacy, Science, and Technology of Targu Mures, Romania
2. Clinical Laboratory, Emergency County Clinical Hospital of Targu Mures, Romania

ABSTRACT

Background: External quality control is a critical component of laboratory medicine, ensuring the accuracy, reliability, and comparability of test results across different laboratories and over time. This study aimed to evaluate the feasability of implementing a split-sample pilot exchange program involving multiple laboratories in Romania.

Methods: The program involved six rounds of a split-sample protocol between 34 public and private Romanian laboratories, including four basic chemistry parameters: glucose, total cholesterol, triglycerides, and HDL-cholesterol. Performance was evaluated with traditional assessment methods, including Z-scores and performance specifications derived from biological variation. Results were analyzed to determine compliance with acceptable limits and performance metrics.

Results: Overall, most laboratories achieved satisfactory results when assessed using mean ±3 standard deviations, particularly for total cholesterol and triglycerides. However, a significant number of laboratories failed to meet optimal Total Allowable Error (TAE) limits for glucose and HDL-cholesterol. While glucose showed satisfactory minimum TAE performance, many laboratories reported values above clinical decision thresholds. Similar trends were observed for HDL-cholesterol, where a notable proportion of laboratories did not meet optimal TAE specifications. These findings suggest that relying solely on Z-scores may be inadequate for performance evaluation.

Conclusions: More stringent performance criteria are warranted to identify areas for improvement, ensuring consistent and reliable laboratory results that can accurately inform clinical decisions across participating laboratories. This could facilitate better identification of corrective actions.

Keywords: ISO 17043, performance evaluation, proficiency testing, Z-score

Received: 6 February 2025; Accepted: 11 March 2025; Published: 27 April 2025.

INTRODUCTION

Proficiency Testing (PT) involves the use of interlaboratory comparisons to evaluate laboratory performance [1]. In the ISO 15189:2022 Standard [2], specific for medical laboratories, External Quality Assessment (EQA) serves as a method to monitor the performance of the examination methods. PT schemes are traditionally organized by commercial providers or by professional societies and many of them are accredited by 17043 Standard [3]. The participants' performance is evaluated by using criteria imposed by the federal law in the United States, CLIA (Clinical Laboratory Improved Amendments) [4]. Other providers use acceptable limits derived from the biological variations or use their own developed limits based on the participant's performances [5,6]. The performance evaluation criteria may be selected by the provider according to the PT scheme objectives [1]. Usually, the EQA schemes that have as the main objective educational purposes, set tighter limits than regulatory programs [7]. There are a few methods for target value assignment and for performance evaluation depending on the type of data obtained and the Z-score is the most commonly used [8]. Alternative approaches of the PT schemes are described in the ISO 15189:2022 Standard as an alternative to the commercial schemes and may be used, in some circumstances, for evaluating laboratories that participate in sample exchange programs [1]. Some of the commercial providers use a non-human serum-based product or spike the samples to reach a certain concentration and this may lead to non-commutability of the material prepared. Using non-commutable samples may result in the impossibility of evaluating equivalence between different laboratories [9].

* Corresponding author: Diana Ionela Handolescu, Clinical Laboratory, Emergency County Clinical Hospital of Targu Mures, Romania. E-mail: dianaionela23@gmail.com

The study aimed to evaluate the performance of the participants in a split-sample protocol using traditional assessment methods (Z score) and compare it to performance specifications derived from the biological variation [10] in four basic chemistry parameters with clinical decision thresholds important for diabetes mellitus diagnosis and cardiovascular risk stratification in the general population.

METHODS

The exchange program study was conducted between September 2023 and March 2024. Sample exchange was organized in six rounds. For each round, a human serum-based sample was prepared, aliquoted and distributed in 34 public and private Romanian participating laboratories (split-sample protocol). All participating laboratories are accredited according to ISO 15189 Standard. From each sample glucose, total cholesterol, triglycerides (TGL) and HDL-cholesterol were measured by each participating laboratory using one or more analyzers. Results were reported via a Google Form. Reported results were then statistically assessed. For each round, the mean value and Standard Deviation (SD) were calculated from the results reported by the participants. Outliers, if present, were identified with the Tukey test and removed from the final data sample and the participants were notified about the excluded outliers. The target value for each round was the consensus value obtained from the participant's results and acceptable limits were +/- 3SD from the mean value. The acceptable limits were also assessed based on the Total Allowable Error (TAE) limits derived from biological variation (optimum and minimum values). The CV% obtained from the reported results within each round were also compared with the CVi% for each parameter and minimum CVa% derived from biological variation (BV). The study had been approved by the institutional ethics committee.

RESULTS

For glucose, in five out of the six rounds, a considerable number of laboratories reported results that were outside the acceptable limits calculated as the optimal TAE derived from the BV. When evaluated according to the limits calculated as +/-3DS, only 2/54 results in round 2 were outside these limits. Detailed results are shown in Table 1.

For Total Cholesterol, in all six rounds, all participants had results within acceptable limits calculated as +/-3DS and minimum TAE. In round five, all participants were within acceptable limits derived from optimal TAE. Results are shown in Table 2.

Results for TGLs are shown in Table 3. Except for round 4 in which 2/47 participants did not meet the acceptable limits derived from optimal TAE, all participants had results within the minimum TAE and +/-3DS.

Table 1. Serum glucose mean values and limits

Round	Mean (+/-3 SD)	3.1% TAE (Optimal)	9.2% TAE (Minimum)	Outside Optimal TAE
R1	110.59 (100.12-121.06)	107.17-114.01	100.42-120.76	14/37
R2	107.22 (99.48-114.96)	103.9-110.44	97.36-117.08	12/54
R3	114.07 (105.28-122.86)	110.54-117.6	103.58-124.56	12/56
R5	123.81 (111.18-136.44)	119.98-127.64	112.42-135.2	19/50
R6	125.78 (113.15-138.41)	121.89-129.67	114.21-137.35	19/49
R4	62.79 (55.76-69.81)	60.85-64.73	57.02-68.56	1/50

All values are expressed in mg/dL. TAE – Total Allowable Error. Outside Optimal TAE represents the number of participants that had results above Optimal TAE calculated at the mean value obtained in the round out of the total number of reported results.

Table 2. Total cholesterol mean values and limits

Round	Mean (+/-3SD)	4.2% TAE (Optimal)	12.5% TAE (Minimum)	Outside Optimal TAE
R1	132.23 (118.7-144.7)	126.68-137.78	115.71-148.75	5/34
R2	203.09 (182.9-223.28)	194.57-211.61	177.71-228.47	5/49
R3	143.65 (133.6-153.64)	137.62-149.68	125.7-161.6	1/44
R4	74.77 (65.83-84.71)	71.63-77.91	65.43-84.11	13/46
R5	134.78 (126.53-143.03)	129.12-140.44	117.94-151.62	0/42
R6	155.08 (140.51-169.5)	148.57-161.59	135.7-174.46	4/41

All values are expressed in mg/dL. TAE – Total Allowable Error. Outside Optimal TAE represents the number of participants that had results above Optimal TAE calculated at the mean value obtained in the round out of the total number of reported results.

Table 3. Triglyceride mean values and limits

Round	Mean (+/-3SD)	12.9 % TAE (Optimal)	38.7 % TAE (Minimum)	Outside Optimal TAE
R1	130.1 (111.26-148.94)	113.32-146.88	79.76-180.44	0/34
R2	160.73 (138.05-183.41)	140-180	98.53-222.93	0/47
R3	122.20 (114.1-130.3)	106.44-137.96	74.91-169.49	0/23
R4	60.90 (48.69-73.11)	53.05-68.75	37.34-84.46	2/47
R5	122.97 (107.25-138.69)	107.11-138.83	75.39-170.55	0/44
R6	154 (135.79-172.21)	134.14-173.86	94.41-213.59	0/41

All values are expressed in mg/dL. TAE – Total Allowable Error. Outside Optimal TAE represents the number of participants that had results above Optimal TAE calculated at the mean value obtained in the round out of the total number of reported results.

For HDL-C, although all participants met the requirements derived from +/-3DS and minimum TAE, a significant number did not meet the requirements derived from optimal TAE (Table 4).

In all six rounds, the CV% obtained within each group for glucose, Total Cholesterol and TGL was lower than the median CV estimate (CVi%) of intra-individual BV. For HDL-C in 2/6 rounds, the CV% was higher than CVi=5.7%. Compared to minimum analytical (CVa) requirements derived from BV the results were: in 3/6 rounds for glucose the CV%>minimum CVa% and in one round the obtained CV%> minimum CVa% for Total Cholesterol. In all six rounds, the TGL's CV<CVa% and the HDL-C's-CV>CVa%. Results are shown in Table 5.

The mean values and distribution of the reported results in the sample exchange program are shown in Figure 1.

Table 4. HDL-cholesterol mean values and limits

Round	Mean (+/-3SD)	5 % TAE (Optimal)	14.9 % TAE (Minimum)	Outside Optimal TAE
R1	34.2 (29.16-39.24)	32.49-35.91	29.11-39.29	12/32
R2	43.01 (34.76-51.26)	40.86-45.16	36.61-49.41	19/45
R3	39.57 (33.39-45.75)	37.6-41.54	33.68-45.56	7/46
R4	21.78 (16.68-26.88)	20.08-23.48	18.54-25.02	15/46
R5	33.2 (27.77-38.36)	31.54-34.88	28.26-38.14	15/42
R6	36.53 (30.26-42.8)	34.71-38.53	31.09-41.97	12/42

All values are expressed in mg/dL. TAE – Total Allowable Error. Outside Optimal TAE represents the number of participants that had results above Optimal TAE calculated at the mean value obtained in the round out of the total number of reported results.

Table 5. Coefficients of variation (CV%)

	Glucose	Total cholesterol	Triglycerides	HDL-cholesterol
R1	3.16	3.40	4.83	4.92
R2	2.41	3.32	4.71	**6.41**
R3	2.57	2.32	2.21	5.22
R4	**3.41**	**3.99**	6.69	7.83
R5	**3.55**	2.05	4.26	5.47
R6	**3.73**	3.13	3.95	**5.73**
CVi	4.6	5.2	19.7	5.7
CVa (minimum)	3.4	3.9	14.8	4.3

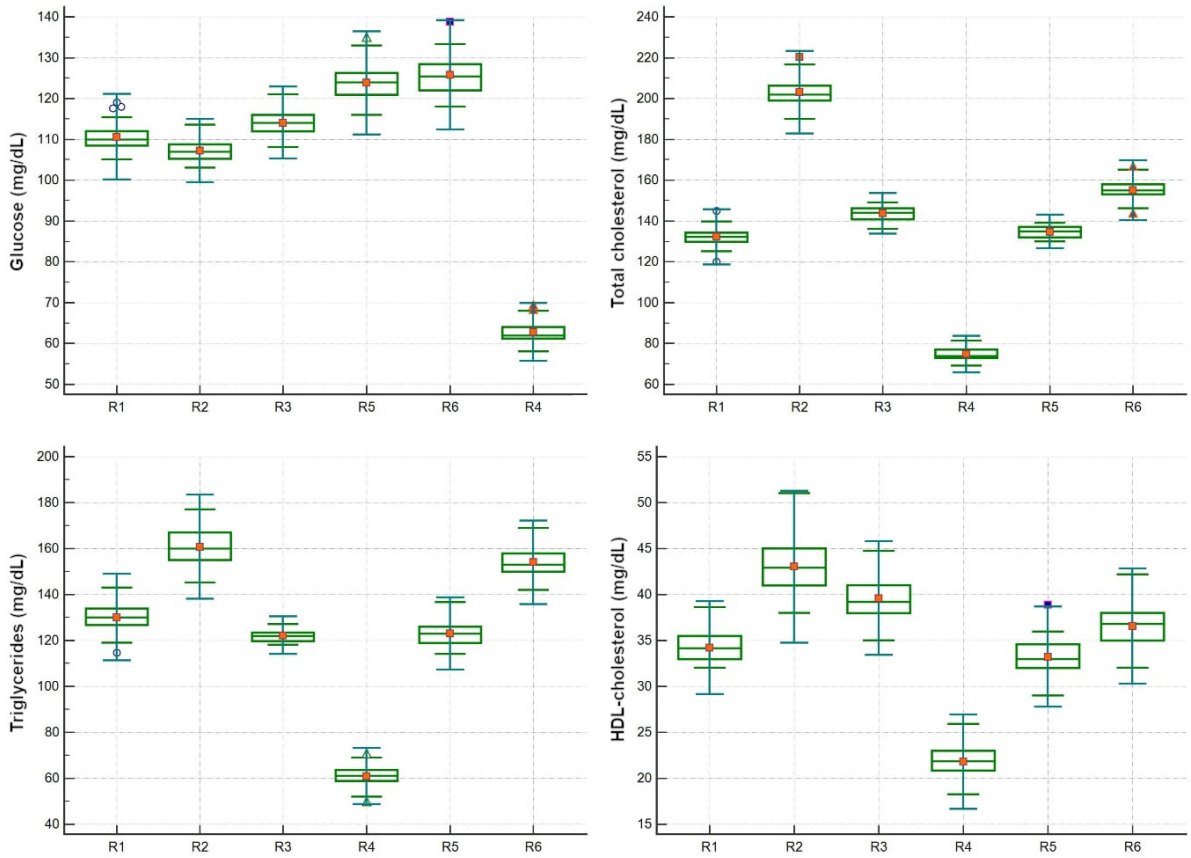

Figure 1. Mean values and distribution of the reported results in the sample exchange program

DISCUSSION

Key findings

With very few exceptions, all laboratories had satisfactory results if the assessment of the performance was made with the mean +/-3DS (Z score) and minimum requirements of TAE based on the biological variations. However, when compared with the optimal TAE requirements, in glucose and HDL-C, a large number of the results did not meet these limits.

Glucose

Although all participants met the minimum TAE derived from BV and the Z scores were <3, in five rounds the limits based on optimal TAE were not met by all participants. The CV for glucose obtained in each round was lower than the CVi% but in half the rounds, when CV% was compared with the analytical (CVa%) derived from biological variation, these specifications were not met. Although the compared CV is not equal to the CV% obtained in one laboratory, the higher CV suggests that the results obtained in different laboratories may not be comparable. Since the samples were human-based, the commutability of the samples does not influence the results depending on the method. In round 6, the Target Value was around the Clinical Decision Limit of 126 mg/dL but 23/49 participants obtained values >126 mg/dL thus including the patient in a category that has one criterion for the DM diagnosis [11]. If we compare the results with the CLIA requirements [4], the +/- 6 mg/dL is also not met much like the optimal TAE derived from BV. In some programs, like the French EQA, this limit is even lower (5.5 mg/dL) than the new CLIA goal [5]. In a published report for US laboratories from 2006, a 2-3% PT failure was reported for glucose [12]. If we evaluate the results in our study only based on the Z-score, almost no failure rate was registered in our study. In a recent study from 2023, the glucose measurement results from 58 Norwegian hospital laboratories participating in two different EQA schemes over one year were analyzed. The study highlighted the impact of commutability and value transfer methods on observed biases in glucose measurements. The authors of the study did not publish the Z scores but the median values of the Bias from the target value which was <5% [13].

Unfortunately, the SEKK and French participants' results were not published probably due to confidentiality reasons so a comparison with participants from these PT providers was not made but considering that the limits in these programs are calculated from participant's results, we can assume that most of them have very good performances.

Total cholesterol

The performances obtained in Total Cholesterol were better than in glucose, although the median CVi% estimate is similar to glucose. In the French EQA program, the limits derived from the participants' performances are lower than the CLIA 2025 limits and the limits derived from BV [5]. A limit of 8.5% is set in the SEKK program derived from the performance of the participants [6]. In round 2, the target value was 203.9 mg/dL but 17/49 results were <200 mg/dL thus including the patient in the optimal level for serum cholesterol in individuals with a low-level cardiovascular risk [14]. The failure rate reported in US laboratories was <3% [12] for Total cholesterol. The results obtained in our study are similar if the performance evaluation is made using the Z-score and minimum TAE requirements. A study that analyzed results from an EQA program found that while the majority of laboratories achieved results within the Analytical Performance Specifications, there were notable differences between methods, with some laboratories exhibiting biases up to 0.91 mmol/L at a median cholesterol level of 5.59 mmol/L [15]. Additionally, the Weqas Reference Laboratory assigns target values for lipid EQA materials, including cholesterol, using the NIST Isotope Dilution Gas Chromatography Mass Spectrometry (ID-GCMS) Reference Method. This approach provides a "true" measurement of cholesterol, ensuring accuracy in assessing participant results against the reference method for each distribution [16].

Triglycerides

The acceptable limits for TGL were met in all rounds but these limits are derived from a median CVi estimate of 19.7%. The French limits and the SEKK limits are much lower, set at 15% and 12% respectively and in the new CLIA regulations these limits were lowered to 15% from 25%. In round six, the TV was 154 mg/dL and all participants met the requirements but 8/41 participants reported results <150 mg/dL, the target value for some lipid-lowering therapies [14]. In a study published in Clinical Chemistry in June 2014, assays for the measurement of eight common analytes, including triglycerides, using a panel of 20 fresh-frozen single-donation serum samples were assessed. The study found that while most assays showed excellent peer performance attributes, there were instances where individual assays had biases exceeding the used limits, such as the Beckman Coulter AU triglycerides assay, which exhibited a 5.4% bias. This highlights the need for improvement in assay standardization and the interchangeability of results [17]. Another study published in Clinical Chemistry and Laboratory Medicine in 2017 evaluated the commutability of EQA materials and the accuracy of routine systems for serum

triglyceride measurements. The study analyzed 43 fresh patient specimens and 32 processed materials across 14 routine methods, revealing that while frozen serum samples and swine sera were commutable for all assays, other EQA materials exhibited matrix effects differently across routine methods. The study concluded that matrix effects and calibration biases existed in triglyceride measurements, indicating a need for continued efforts to improve accuracy and comparability [18].

HDL-cholesterol

For HDL-C, some of the results were not between the optimal TAE limits. In 2/6 rounds the CV% obtained was higher than the CVi% median estimate. CLIA has lowered the limit for acceptability from 30% to 20% [4]. The minimum TAE limit is 14.9% and only 2 laboratories did not meet this requirement. In round 2, the TV was 43.01 mg/dL and 33/45 participants reported results <45 mg/dL but 12 reported results >45 mg/dL.

Total error ranged from −26% to +32% for direct LDL-Cholesterol and −20% to +36% for direct HDL-C in a comprehensive study of different assays in dyslipidemia samples [19]. Most discordances in dyslipidemia samples are observed at lower concentration ranges of HDL-C (< 40 mg/dL) [19]. These errors may result in misclassifications for ASCVD risk assessment depending on the type of assay, as observed in accuracy-based external quality assessment surveys of hypertriglyceridemia samples organized across different laboratories [20]. The biases noted in direct HDL-C measurements affect the calculations of LDL-C and non-HDL-C, since HDL-C is used in the calculations [21].

In a study from 2022 that analyzed the accuracy, precision, and trueness of HDL-C measurements across participating laboratories in a four-year trueness-based EQA program in China, approximately 95% of laboratories met the minimum TAE criteria derived from BV but challenges remained in achieving optimal measurement standardization for HDL-C. The study emphasized the need for continued efforts to improve trueness and comparability in HDL-C measurements [22].

The Dutch EQA organizer, the Stichting Kwaliteitsbewaking Medische Laboratoriumdiagnostiek (SKML), conducts accuracy-based EQA schemes for clinical chemistry analytes, including HDL-C. These programs involve the preparation of serum pools, which are value-assigned for HDL-C using CDC Reference Methods. Such initiatives provide valuable data on laboratory performance and highlight the importance of commutable materials in assessing measurement accuracy [21].

Obtaining a Z score of < +/- 3 in PT testing may not be sufficient for evaluating the participating laboratories' performances as shown by Cordeiro et.al [23]. The authors propose a tool for including measurement uncertainty in the performance evaluation that may lead to better identification of the need for corrective actions even if the Z score is <+/-3 [23].

CONCLUSIONS

The evaluation of the participants' performance in the pilot sample-exchange program for basic clinical chemistry showed that almost all laboratories have a satisfactory result if a large interval is used. When small intervals are used for performance evaluation, a large proportion of the laboratories did not meet these requirements. If the results are reported to clinical decision levels, not all participants included the patients in the same clinical category. More stringent requirements are needed for the participants in EQA programs that may help to identify the opportunities for improvement.

ABBREVIATIONS

BV – biological variation

CLIA – Clinical Laboratory Improved Amendments

CV – coefficient of variation

EQA – external quality assessment

PT – proficiency testing

SD – standard deviation

TAE – total allowable error

TGL – triglycerides

ACKNOWLEDGEMENT

This study was funded by the Romanian Association of Laboratory Medicine (AMLR) grant no. 10/29.07.2024.

AUTHORS' CONTRIBUTION

ORO – conceptualization, methodology, investigation, formal analysis, writing (original draft preparation)

DIH – data curation, statistical analysis, investigation, manuscript preparation

MD – supervision, writing (review and editing)

CONFLICT OF INTEREST

None to declare.

Publisher's Note: The Editorial Office stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. ISO 17043:2023 Conformity Assessment-General requirements for the competence of proficiency testing providers (ISO/IEC 17043:2023).
2. ISO 15189: 2022 Medical Laboratories- Requirements for Quality and Competence.
3. Eptis Database for PT Providers. https://www.eptis.bam.de/eptis/ . Accessed January 2025.
4. CLIA 2025 Acceptance Limits for Proficiency Testing. https://westgard.com/clia-a-quality/quality-requirements/2024-clia-requirements.html . Accessed January 2025.
5. French EQA performance specifications. https://westgard.com/clia-a-quality/quality-requirements/french-eqa-goals.html . Accessed January 2025.
6. A selection of SEKK-DMax specifications https://westgard.com/clia-a-quality/quality-requirements/sekk-dmax.html . Accessed January 2025.
7. Buchta C, Marrington R, De la Salle B, Albarède S, Badrick T, Bietenbeck A, et al. Behind the scenes of EQA - characteristics, capabilities, benefits and assets of external quality assessment (EQA): Part I- EQA in general and EQA programs in particular. Clin Chem Lab Med. 2025 aop. DOI: 10.1515/cclm-2024-1289
8. Eurachem Guide: Selection, Use and Interpretation of Proficiency Testing (PT) Schemes, 3rd Edition. 2021.
9. Sandberg S, Fauskanger P, Johansen JV, Keller T, Budd J, Greenberg N, et al. IFCC Working Group on Commutability in Metrological Traceability. Recommendations for Setting a Criterion and Assessing Commutability of Sample Materials Used in External Quality Assessment/Proficiency Testing Schemes. Clin Chem. 2023 2;69(11):1227-37. DOI: 10.1093/clinchem/hvad135
10. EFLM Biological Variation Database: https://biologicalvariation.eu . Accessed January 2025.
11. American Diabetes Association Professional Practice Committee; 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024. Diabetes Care. 2024;47(S1):S20-S42. DOI: 10.2337/dc24-S002
12. Howerton D, Krolak JM, Manasterski A, Handsfield JH. Proficiency Testing Performance in US Laboratories: Results Reported to the Centers for Medicare & Medicaid Services, 1994 Through 2006. Arch Pathol Lab Med. 2010 134(5):751-8. DOI: 10.5858/134.5.751
13. Gidske G, Sandberg S, Fauskanger P, Pelanti J, Tollånes M, Solsvik A, et. al. Aggregated data from the same laboratories participating in two glucose external quality assessment schemes show that commutability and transfers of values to control materials are decisive for the biases found. Clin Chem Lab Med. 2024;62(1):77-84. DOI: 10.1515/cclm-2023-0532
14. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. ESC Scientific Document Group , 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). Eur Heart J. 2020;41(1):111-8.
15. Robbins N, Shepherd S, Graham P. Cholesterol measurement - unexpected results in a commutable EQA program. Pathology. 2022;54(S1):S65-S66. DOI: 10.1016/j.pathol.2021.12.214
16. WEQAS website https://www.weqas.com/services/referencelab/gcms-cholesterol . Accessed January 2025.
17. Hedwig C, Stepman M, Tiikkainen U, Stöckl D, Vesper HW, Edwards SH, et al. on behalf of the participating laboratories. Measurements for 8 Common Analytes in Native Sera Identify Inadequate Standardization among 6 Routine Laboratory Assays. Clin Chem. 2014;60(6):855-63. DOI: 10.1373/clinchem.2013.220376
18. Meng Q, Zhou W, Zhang C, Zeng J, Zhao H, Zhang T, et al. Serum triglyceride measurements: the commutability of reference materials and the accuracy of results. Clin Chem Lab Med. 2017 55(9):1284-90. DOI: 10.1515/cclm-2016-0682
19. Borge G, Langlois MR, Langsted A, Chapman JM, Aakre KM, Baum H et al. Quantifying atherogenic lipoproteins for lipid-lowering strategies: Consensus-based recommendations from EAS and EFLM, Atherosclerosis, 2020 294: 46- 61. DOI: 10.1016/j.atherosclerosis.2019.12.005
20. Miller WG, Myers GL, Sakurabayashi I, Bachmann LM, Caudill SP, Dziekonski a et al. Seven direct methods for measuring HDL and LDL cholesterol compared with ultracentrifugation reference measurement procedures, Clin Chem 2010 56: 977-986. DOI: 10.1373/clinchem.2009.142810
21. Langlois MR, Descamps OS, van der Laarse A, Weykamp C, Baum H, Pulkki K, et al. Clinical impact of direct HDLc and LDLc method bias in hypertriglyceridemia. A simulation study of the EAS-EFLM Collaborative Project Group. Atherosclerosis. 2014;233:83-90. DOI: 10.1016/j.atherosclerosis.2013.12.016
22. Zhou W, Luo W, Yu S, Li H, Wang D, Zhang J, et al. Performance of HDL-C measurements assessed by a 4-year trueness-based EQA/PT program in China. Clin Chem Lab Med. 2022;60(10):1586-97. DOI: 10.1515/cclm-2020-0658
23. Cordeiro F, Emons H, Robouch P. Is the z score sufficient to assess participants' performance in proficiency testing? The hidden corrective action. Accreditation and Quality Assurance. 2022;27:145-53. DOI: 10.1007/s00769-022-01496-w