International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

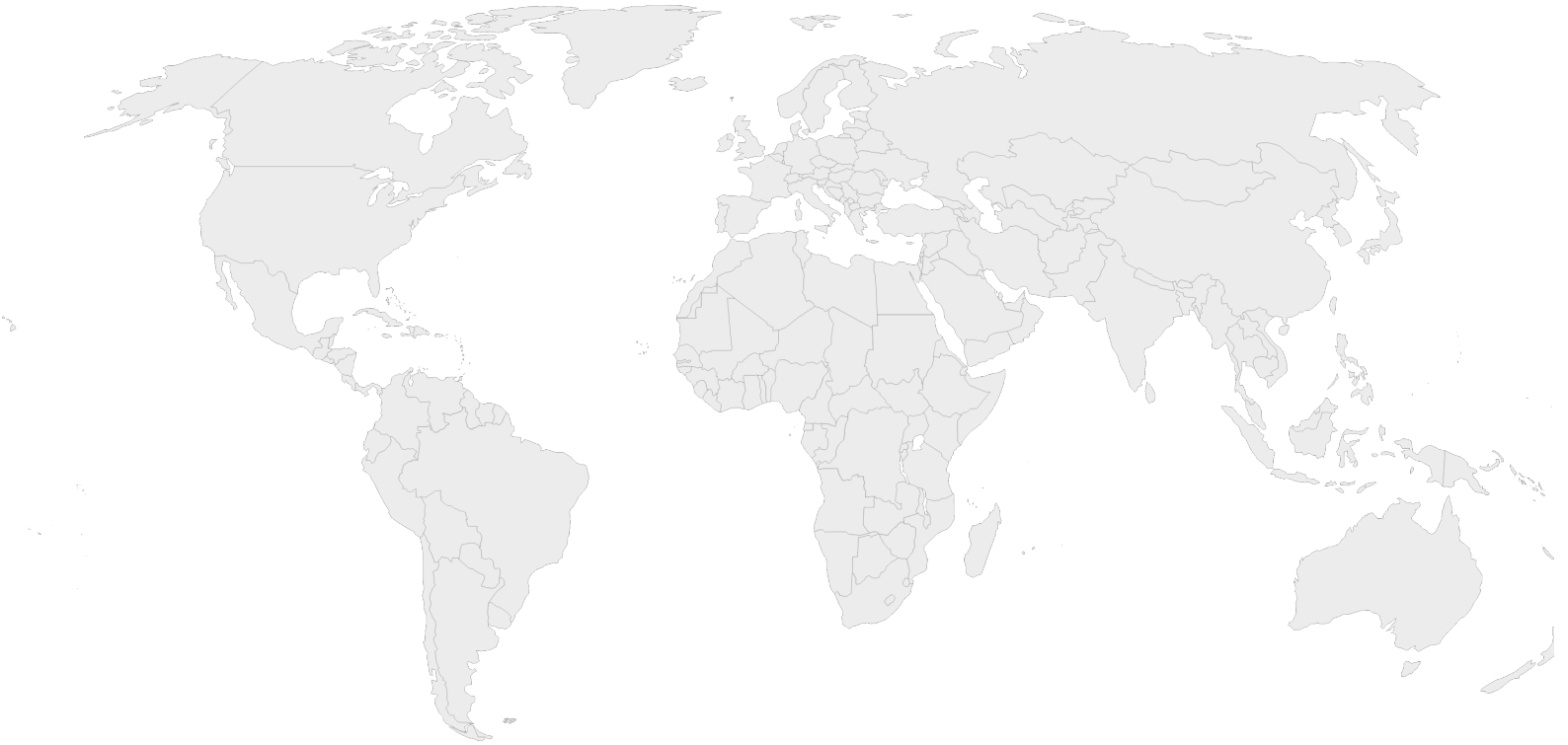
Communications and Publications Division (CPD) of the IFCC

August 2026
Volume 37
Number 4
ISSN 1650-3414

Co-Editor-in-Chief:
Qing H. Meng, MD, PhD, DABCC, FADLM
Department of Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, USA

Co-Editor-in-Chief:
Kannan Vaidyanathan, MD, FRCP (Edin.) (UK), FFSc (Research) (RCPA) (Australia), FACBI
Department of Biochemistry
Believers Church Medical College Hospital
Thiruvalla, Kerala, India

E-mail: ejifcc@ifcc.org



In this Issue

In this Issue

In this Issue

Research Article

Ethical Challenges in Point-of-Care Testing for Cardiac Biomarkers: Balancing Rapid Diagnosis with Clinical Responsibility

Sibtain Ahmed[1] , Melissa M. Budelier[2, 3] , Nafija Serdarevic[4], Joesph Wiencek[5*]

[1]*Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*
[2]*TriCore Reference Laboratories, Albuquerque, NM, USA*
[3]*Department of Pathology, University of New Mexico School of Medicine, Albuquerque, NM, USA*
[4]*Department of Clinical Biochemistry and Laboratory Medicine, University Clinical Center, Sarajevo, Faculty of Health Studies, University of Sarajevo, Sarajevo, Bosnia and Herzegovina*
[5]*Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA*

Article Info

Corresponding Author:
Joesph Wiencek
Associate Professor of Pathology, Microbiology and Immunology
Vanderbilt University School of Medicine
Nashville TN, USA
E-mail: joe.wiencek@vumc.org

Keywords

Point-of-care testing, Cardiac biomarkers, Ethics

Abstract

The rapid expansion of point of care testing (POCT) has significantly transformed clinical decision-making, particularly in acute care settings. Cardiac biomarkers such as troponins (conventional and high sensitivity) and natriuretic peptides are increasingly being deployed at the bedside to expedite diagnosis and guide timely management [1]. While the clinical advantages of POCT are well recognized, its widespread implementation raises important ethical considerations that remain insufficiently addressed [2]. We aim to project that the rapid expansion of cardiac biomarker POCT has outpaced the development of practical ethical guidance for its implementation. In particular, challenges related to analytical reliability, accountability, quality oversight, equitable access, patient autonomy, and environmental sustainability require a structured ethical framework to support safe and clinically responsible use of POCT in cardiac care. One of the primary ethical concerns relates to the trade-off between speed and analytical accuracy. Although POCT offers faster turnaround times, variability in analytical performance compared to central laboratory assays may compromise diagnostic certainty [3]. This is particularly critical in the context of cardiac biomarkers, where even minor inaccuracies can lead to misclassification of myocardial infarction. In cardiac care, such errors carry particularly significant ethical implications because troponin results may directly influence urgent decisions regarding admission, invasive procedures, discharge, or withholding of treatment. This underscores the necessity of an ethical framework to ensure rapid testing does not compromise diagnostic integrity and patient safety.

Unlike conventional laboratory testing, POCT is often performed in environments where critical factors such as temperature, humidity, time to analysis, and oxygen exposure are less controlled and may fluctuate significantly, thereby affecting assay performance [4]. Furthermore, the pre-analytical phase represents a major source of error in POCT. The use of whole blood samples, particularly in capillary (fingerstick) testing, increases susceptibility to hemolysis, while improper specimen handling can introduce errors such as air bubbles, micro clots, or gross clotting [5, 6].

Another important issue is the question of accountability and diagnostic testing acumen. POCT is often performed by non-laboratory personnel, which may lead to variability in test execution and interpretation but also in the level of technical competence compared to trained laboratory professionals [2]. While non-laboratory personnel may be willing to accept responsibility for performing and interpreting tests, this does not necessarily ensure adequate proficiency or adherence to standardized procedures. This concern is particularly relevant for cardiac biomarkers, where high-sensitivity assays are increasingly integrated into established rule-in and rule-out algorithms developed and verified primarily in central laboratory settings in collaboration with specialties such as emergency medicine and cardiology. However, the direct applicability of these interpretive pathways in POCT environments remains incompletely established because POCT testing may be more vulnerable to preanalytical variables, operator-dependent factors, and analytical variability. Although recent studies have begun evaluating point-of-care high-sensitivity troponin assays and accelerated diagnostic pathways in emergency care settings, evidence remains comparatively limited when compared with well-established central laboratory protocols [7]. Further multicenter validation studies are therefore needed before these algorithms can be universally adopted across diverse POCT settings. From an ethical and clinical perspective, cardiac biomarker POCT may be most appropriate in settings where rapid triage and timely decision-making are critical, particularly when central laboratory access is delayed or unavailable. However, for complex clinical scenarios requiring highly standardized serial testing, comprehensive diagnostic interpretation, or confirmatory evaluation, central laboratory testing may remain the preferred approach. An ethical framework is therefore essential to clearly define roles, responsibilities, and oversight while also ensuring appropriate training and competency assessment to safeguard the quality of patient care.

Quality assurance represents an additional challenge. POCT environments may lack standardized oversight compared to central laboratories, raising concerns related to patient safety and non-maleficence [8]. This highlights the need for ethical and operational guidance including the implementation of standardized algorithms, training programs, regular competency assessments, and clearly defined governance frameworks with laboratory oversight. In addition, harmonization of test results between POCT devices and central laboratory assays remains a significant challenge, particularly in ensuring comparability and consistency of results for patients managed within the same healthcare setting [8]. In the interim, it is imperative these two types of measurements be clearly delineated from clinical interpretations to avoid harm.

The ease of access to POCT devices may also contribute to overutilization of cardiac biomarker testing, increasing the chance for false results and unnecessary interventions [10]. However, these devices are also crucial in improving access to timely diagnostics, particularly for patients in resource-limited and rural settings where centralized laboratory services may be unavailable or delayed. By enabling rapid decision-making at the point of care, POCT can facilitate early diagnosis, triage, and management of acute cardiac conditions, thereby potentially improving clinical outcomes. Nevertheless, disparities in implementation and resource allocation may exacerbate existing healthcare inequities if access to POCT technologies is uneven [11]. These ethical challenges are summarized in Table 1. An ethical blueprint is thus crucial to promote appropriate utilization while ensuring equitable access to POCT services.

From cardiac biomarkers perspective, due to their role in high-stakes decision-making ethical concerns become paramount. Addressing these challenges requires strengthening training, ensuring laboratory oversight, implementing robust quality systems, and developing ethical frameworks.

While POCT has clear advantages in improving timely detection of acute cardiac events, it raises important questions about which high-risk conditions are suitable for decentralized testing. It also highlights concerns about how complex, time-sensitive clinical information is interpreted and communicated when testing is performed by non-laboratory personnel outside the central laboratory, even within supervised care settings. At the same time, as patients increasingly access personal health data through electronic records and wearable technologies, an ethical dilemma emerges between promoting autonomy and ensuring beneficence and non-maleficence, underscoring the need for clear safeguards and clinical oversight in the expanding use of cardiac POCT. Figure 1 illustrates the balance between clinical benefits and ethical risks associated with POCT for cardiac biomarkers, emphasizing the need for quality assurance and governance.

Moreover, the increasing use of POCT has also brought attention to an important ethical concern related to environmental sustainability. Many POCT devices rely heavily on single-use plastics, including cartridges, test strips, lancets, and packaging materials. As the use of POCT continues to expand, particularly in high-volume areas such as emergency cardiac care, the amount of biomedical plastic waste generated is also increasing. Previous reports have highlighted that disposable diagnostic materials contribute significantly to healthcare-related plastic waste, which may create environmental and public health challenges, especially in

settings where waste management systems are limited [12]. Although POCT provides substantial clinical benefits through rapid diagnosis and timely patient management, these advantages should be balanced against their broader environmental impact. From an ethical perspective, this relates to the principles of non-maleficence and justice, as healthcare practices should aim to avoid unintended harm to both communities and the environment. Therefore, alongside improving patient care, there is also a need to encourage sustainable disposal practices and support the development of more environmentally responsible diagnostic technologies.

In conclusion, the expanding role of POCT, particularly for high-risk applications such as cardiac biomarkers, necessitates a robust ethical framework to guide its implementation. As also emphasized by the Canadian Society of Clinical Chemists, POCT should be conducted in a manner that prioritizes effective patient care while judiciously utilizing available resources, with financial or personal incentives such as making workflows easier for clinicians, the push to provide faster results, meeting performance targets, or possible influence from industry, not serving as primary drivers [13]. Equally important are the principles of patient autonomy and confidentiality, particularly as POCT becomes increasingly integrated with digital health technologies, wearable devices, and personal health records. In these evolving care models, patients should be adequately informed not only about the purpose and limitations of testing, but also about how their results may be stored, accessed, shared, or integrated across healthcare systems. This is especially relevant as rapid connectivity, and remote monitoring may increase the risk of unauthorized access or unintended disclosure of sensitive health information. Therefore, informed consent processes and robust data protection safeguards are essential to preserve patient trust, maintain confidentiality, and ensure ethical use of POCT-generated data. In this rapidly evolving landscape, the vital and timely role of the IFCC Task Force on Ethics is particularly significant in providing global guidance and standard setting to address emerging ethical challenges in POCT. An ethical framework for POCT in cardiac care should include key elements such as quality assurance processes, protection of patient confidentiality, equitable access to testing, and consideration of environmental sustainability. Such measures may help ensure that the expanding use of POCT remains aligned with patient safety and ethical clinical practice. As access to rapid, decentralized cardiac testing continues to expand, maintaining a balance between accessibility, clinical oversight, and ethical responsibility will be essential to ensure safe, equitable, and patient centered care.

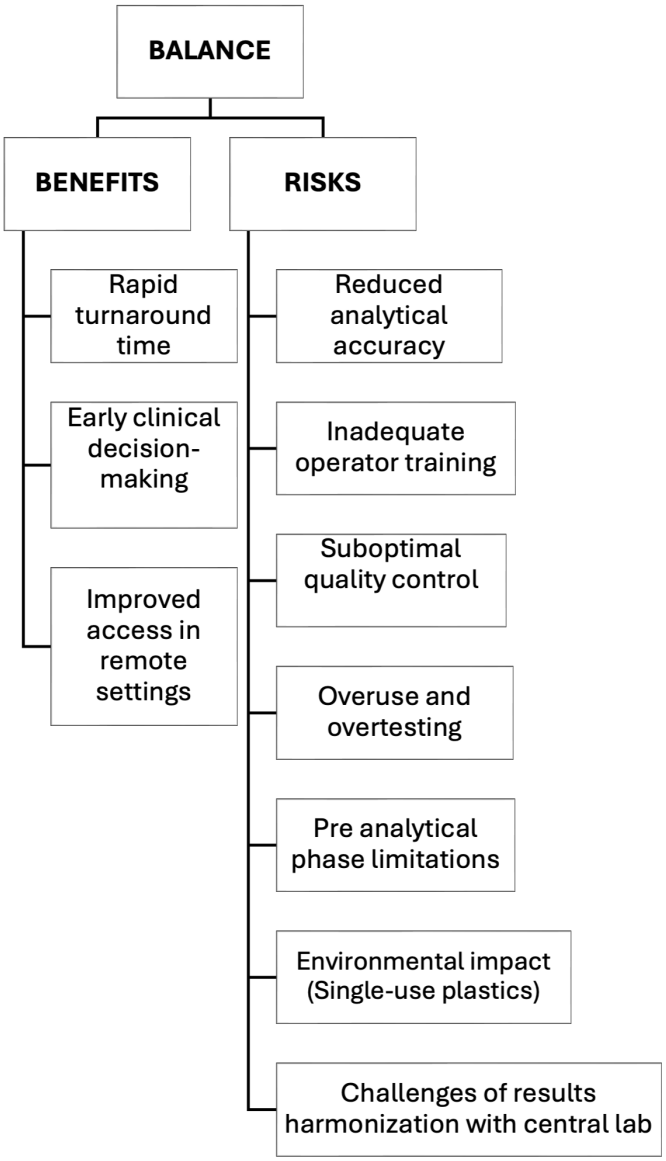
Table 1: Ethical considerations in POCT for cardiac biomarkers and their clinical implications.

Ethical Principle	Ethical Considerations	Description	Clinical Impact
Autonomy (Clinical decision-making responsibility)	**Professional Accountability and Governance** **Informed consent and data confidentiality**	Testing is often performed by non-laboratory personnel	Risk of misinterpretation and inappropriate clinical decisions (e.g., non-lab operators interpreting delta troponins, both absolute and percentage deltas) Patients may not fully understand how POCT-generated data are stored, shared, or integrated with electronic health systems
Beneficence (Act in patient's best interest)	**Improved Access**	POCT expands access to diagnostics, especially in remote or resource-limited settings, enabling earlier identification and triage of acute conditions when central laboratory testing is not readily available.	Timely clinical decision-making, improved patient outcomes, and more appropriate triage of MI cases (e.g., identifying patients requiring urgent versus non-urgent care).
Equity (Fair distribution of care)	**Resource Disparity**	Use in low-resource settings without adequate infrastructure or quality systems; however, POCT can significantly improve access to diagnostic testing in rural or underserved areas where central laboratory services are unavailable.	Potential variability in test quality and accuracy; however, improved access may enable earlier triage and clinical decision-making, particularly for patients distant from healthcare facilities.

Ethical Principle	Ethical Considerations	Description	Clinical Impact
Non-maleficence (Do no harm)	**Accuracy vs Speed**	POCT may have lower sensitivity compared to central lab testing, though newer high-sensitivity assays are improving	Missed or delayed diagnosis of MI (e.g., false negative MI)
	Quality Assurance	Variability due to preanalytical errors and operator dependence; inconsistent QC/ EQA practices	Variable reliability of results, potential patient harm
	Environmental Impact (Single-use plastics)	POCT generates single-use plastic waste (e.g., cartridges, strips), raising environmental concerns	Environmental contamination, long-term public health risks, and ethical concerns regarding sustainability and responsible healthcare delivery
Justice (Fair use of resources)	**Over-testing**	Easy availability may lead to unnecessary or excessive testing (e.g., troponin overuse)	False positives, unnecessary admissions, increased healthcare costs

POCT, Point of Care Testing; QC, Quality Control; EQA, External Quality Assessment; MI, Myocardial Infarction

Figure 1: Ethical balance in POCT for cardiac biomarkers.



Disclosures of Interest

RW: Research support Bio-Rad Laboratories. Honoraria: Bio-Rad Laboratories. Travel Support: IFCC, ADLM, EFLM, ASCP. Leadership Roles: IFCC Chair of the Task Force on Ethics, ASCP Chair Clinical Chemistry Special Interest Group; Editorial Board, Journal of Applied Laboratory Medicine. Consultant: Legion Diagnostics. MMB: Advisory Board: Roche Diagnostics, Greiner Bio-One; Travel Support and Lecture Fees: Roche Diagnostics, Bio-Rad, ADLM, MSACL. Editorial Board: Associate Editor, Journal of Mass Spectrometry and Advances in the Clinical Lab and Archives of Pathology; Receiving Editor, Clinica Chemica Acta. Leadership: Chair, CLN Board of Editors. SA: speaker/consultancy services agreement Roche Diagnostics.

Ethical Approval

Not required as study does not involve human subjects or data.

Funding

None.

Data Availability

Not applicable.

Author Statement

S.A. contributed to major writing, preparation of tables and figures, and literature search. M.M.B. contributed to writing, table preparation, and review. N.S. contributed to writing and review. J.R.W. contributed to conceptualization, writing, and review. All authors reviewed and approved the final manuscript

References

1. Ahmed S, Jafri R, Mahmood A. POCT in cardiovascular operating room: a game changer? EJIFCC. 2025;36, 3: 222–224. PMID: 41098222
2. Nichols JH. Point of care testing. Clin Lab Med. 2007;27(4):893–908. doi:10.1016/j.cll.2007.07.003
3. Apple FS, Collinson PO. Analytical characteristics of high-sensitivity cardiac troponin assays. Clin Chem. 2012;58(1):54–61. doi:10.1373/clinchem.2011.165795
4. Kazmierczak SC, Morosyuk S, Rajkumar R. Evaluation of Preanalytical Point-of-Care Testing Errors and Their Impact on Productivity in the Emergency Department in the United States. J Appl Lab Med. 2022;7(3):650-660. doi: https://doi.org/10.1093/jalm/jfab158
5. O'Hara M, Wheatley EG, Kazmierczak SC. The Impact of Undetected In Vitro Hemolysis or Sample Contamination on Patient Care and Outcomes in Point-of-Care Testing: A Retrospective Study. J Appl Lab Med. 2020;5(2):332-341. doi: https://doi.org/10.1093/jalm/jfz020
6. Krasowski MD. Educational Case: Hemolysis and Lipemia Interference With Laboratory Testing. Acad Pathol. 2019;6:2374289519888754. doi: https://doi.org/10.1177/2374289519888754
7. Cullen L, Collinson PO, Giannitsis E. Point-of-care testing with high-sensitivity cardiac troponin assays: the challenges and opportunities. Emerg Med J. 2022;39(11):861-866. PMID: 35017187. doi: 10.1136/emermed-2021-211907
8. Point-of-care testing (POCT)— Requirements for quality and competence (Available from: https://www.iso.org/obp/ui/#iso:std:iso:22870:ed-2:v1:en) (Accessed on: 09-05-2026)
9. Plebani M, Nichols JH, Luppa PB, Greene D, Sciacovelli L, Shaw J, Khan AI, Carraro P, Freckmann G, Dimech W, Zaninotto M, Spannagl M, Huggett J, Kost GJ, Trenti T, Padoan A, Thomas A, Banfi G, Lippi G. Point-of-care testing: state-of-the art and perspectives. Clin Chem Lab Med. 2024;63(1):35-51. doi: 10.1515/cclm-2024-0675. PMID: 38880779.
10. Body R, Carlton E. Understanding cardiac troponin testing. BMJ. 2018;360:k902. doi:10.1136/bmj.k902
11. Pai NP, Vadnais C, Denkinger C, Engel N, Pai M. Point-of-care testing for infectious diseases: diversity, complexity, and barriers. PLoS Med. 2012;9(9):e1001306. doi:10.1371/journal.pmed.1001306
12. Street A, Lie AK, Strasser BJ, Greene JA. Plastics in health care: rethinking medical device innovation, use, and disposal for sustainability. Lancet (London, England). 2025:S0140-6736. doi: https://doi.org/10.1016/s0140-6736(25)01099-2
13. Yip PM, Venner AA, Shea J, Fuezery A, Huang Y, Massicotte L, et al. Point-of-care testing: a position statement from the Canadian Society of Clinical Chemists. Clin Biochem. 2018;53:156–159. doi:10.1016/j.clinbiochem.2017.12.012

Letter to the editor

Rethinking the limit of detection for serum ethanol: evidence from a clinical laboratory

José Arnau Pulido[1], Alejandra Pérez[1] Yolanda Fernández-Cladera[1], María Doladé[1], Susana Malumbres[1], Carla Fernández-Prendes[1*]

[1]*Laboratory Medicine Department, Hospital Universitari Germans Trias i Pujol, Badalona, Spain*

Article Info

[*]***Corresponding Author:***
Carla Fernández-Prendes
Laboratory Medicine Department, Hospital Universitari Germans Trias i Pujol, Badalona, Spain
E-mail: cfernandezp.germanstrias@gencat.cat

Keywords

ethanol, detection, alcohol

Abstract

Ethanol is a widely consumed psychoactive substance. Serum concentrations above 100 mg/dL can cause cognitive impairment and confusion, while levels exceeding 300 mg/dL may be fatal. Thus, rapid and accurate determination of blood ethanol concentration is crucial for diagnosing and managing acute intoxication. Ethanol measurement also has important legal implications, such as in traffic accidents or criminal proceedings.

Gas Chromatography-Mass Spectrometry (GC-MS) is the reference method for ethanol determination due to its accuracy and specificity. In routine clinical laboratories, however, ethanol quantification is usually performed using enzymatic spectrophotometric assays. Alcohol dehydrogenase (ADH) oxidizes ethanol to acetaldehyde with NAD as coenzyme, which is simultaneously reduced to NADH. The increase in NADH absorbance at 340 nm allows indirect quantification of ethanol.

Although analytical characterization and method optimization are the manufacturers' responsibilities, literature indicates these processes are not standardized, and specific procedures are generally not disclosed. Consequently, clinical laboratories may need to verify these values, particularly during implementation of new methods, internal validation, or when analytical results conflict with clinical presentation.

We report four cases in our laboratory with positive ethanol results of 11–14 mg/dL that, after reprocessing, were reported as "not detected". This highlights the need to determine the smallest analyte quantity detectable with certainty, commonly expressed as the limit of detection (LoD).

In this study, we evaluated the LoD of the Emit II Plus Ethyl Alcohol assay (Siemens Healthineers, Dublin, Ireland) in human serum using the AU[5800] analyzer (Beckman Coulter, Brea, CA, USA), following CLSI EP[17]-A[2] recommendations for LoD verification. According to CLSI EP[17]-A[2], LoD is defined as the measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a measurand in a material is β, given a probability α of falsely claiming its presence, and LoB is the highest measurement result that is likely to be observed (with a stated probability α) for a blank sample. LoQ is defined as the lowest amount of a measurand in a material that can be quantitatively determined with stated accuracy, expressed as total error or as independent requirements for bias and precision, under stated experimental conditions. Because our study did not assess predefined accuracy goals for quantitative determination at low ethanol concentrations, it should be interpreted as a local LoD verification study rather than a LoQ evaluation. In this context, the kit insert reports a quantification range of 10–600 mg/dL, with a LoD < 10 mg/dL and reported precision of 0.8%–4.1%. The study was conducted in Spain between April and May 2025, using two instrument systems and two reagent lots to ensure adequate measurements.

An initial estimated LoB was obtained from 20 replicate measurements of ethanol-free serum, defined as the highest observed value (4 mg/dL). This preliminary LoB was then used to guide the selection of low-level samples, targeting concentrations approximately within 1–5 times the LoB, resulting in the preparation of five serum samples (4–20 mg/dL). The experimental LoB was then determined from five ethanol-free serum samples analyzed in four replicates per day over three consecutive days on each instrument. The 95th percentile (α = 0.05) was calculated using the rank-position formula (Rank Position = 0.5 + B·0.95, where B is the total number of replicates). In both analysers, interpolation between adjacent values was required, yielding LoB values of 3 mg/dL for instrument 1 and 9 mg/dL for instrument 2. To minimize false positives, the higher LoB (9 mg/dL) was used for subsequent calculations.

Each low-level serum sample was analyzed in four replicates per day over three days on each instrument, and the LoD for each system was calculated using the formula:

LoD = LoB + cp × SDL,

where SDL represents the pooled standard deviation of the low-level samples, calculated using the following formula:

$$SD_L = \sqrt{\frac{\sum_{i=1}^{J}(n_i - 1)\,SD_i^2}{\sum_{i=1}^{J}(n_i - 1)}}$$

SDi = standard deviation of results for the i-th low-level sample

ni = number of results for the i-th low-level sample

J = number of low-level samples

The correction factor cp is defined as:

$$cp = \frac{1.645}{1 - \frac{1}{4 \times (L - J)}}$$

L denotes the total number of low-level results per instrument/lot (L = 60), J the number of low-level samples (J = 5).

The LoD was 12 mg/dL in instrument system 1 and 14 mg/dL in instrument system 2, calculated using the SDL values of 1.68 and 3.04 mg/dL, respectively.

The discrepancy between the manufacturer-reported LoD (10 mg/dL) and the experimentally determined values (12 and 14 mg/dL) suggests that results may vary with instrument system, reagent lot, and calibration. Therefore, reagent manufacturers should conduct performance studies that incorporate analytical variability, in accordance with CLSI EP[17]-A[2] (6). Yet, standardized protocols are often lacking, and procedures to establish analytical characterization are generally undisclosed. Literature addressing these assay limitations also remains scarce.

The results highlight the importance of evaluating the LoD of enzymatic ethanol assays to minimize false positives. To ensure a conservative threshold, the highest experimentally determined LoD should be adopted. This is particularly relevant given the ethical implications of false-positive results, especially in vulnerable populations such as children or pregnant women. Adopting a higher, experimentally verified LoD does not modify established clinical or legal decision limits but improves the reliability of reporting at very low concentrations by reducing false-positive classifications below the locally verified detection capability of the assay.

In conclusion, for clinical laboratories using the Emit II Plus Ethyl Alcohol assay (Siemens Healthineers, Dublin, Ireland) on the AU[5800] analyzer (Beckman Coulter Inc., Brea, CA, USA), LoD verification is recommended. Consequently, results below the locally verified LoD should not be reported as positive quantitative ethanol results. When discrepancies are clinically or legally significant, the presence or absence of ethanol should be confirmed using a reference method such as GC-MS.

Acknowledgments

None.

Author contributions

Conceptualization: Fernández-Prendes Carla, Fernández Prendes C; José Arnau Pulido, Pulido JA

Methodology: Fernández-Prendes Carla, Fernández Prendes C; José Arnau Pulido, Pulido JA

Investigation: Fernández-Prendes Carla, Fernández Prendes C; José Arnau Pulido, Pulido JA

Visualization: Malumbres Susana, Malumbres S; Maria Doladé, Doladé M; Fernández-Cladera Yolanda, Fernández-Cladera Y

Supervision: Fernández-Prendes Carla, Fernández Prendes C

Writing – original draft: Fernández-Prendes Carla, Fernández Prendes C; José Arnau Pulido, Pulido JA
Writing – review & editing: Fernández-Prendes Carla, Fernández Prendes C

Conflicts of interest
None declared.

Research funding
None.

References

1. Tian, G., Zhang, XQ., Zhu, MS. et al. Quantification of ethanol in plasma by electrochemical detection with an unmodified screen printed carbon electrode. Sci Rep 6, 23569 (2016).
2. Xiao HT, He L, Tong RS, Yu JY, Chen L, Zou J, Li JQ, Bian Y, Zhang Y. Rapid and sensitive headspace gas chromatography-mass spectrometry method for the analysis of ethanol in the whole blood. J Clin Lab Anal. 2014;28(5):386-390
3. Kristoffersen L, Skuterud B, Larssen BR, Skurtveit S, Smith-Kielland A. Fast quantification of ethanol in whole blood specimens by the enzymatic alcohol dehydrogenase method. Optimization by experimental design. J Anal Toxicol. 2005;29(1):66-70.
4. Gegenschatz SA, Chiappini FA, Teglia CM, Muñoz de la Peña A, Goicoechea HC. Binding the gap between experiments, statistics, and method comparison: A tutorial for computing limits of detection and quantification in univariate calibration for complex samples. Anal Chim Acta. 2022;1209:339342. doi: 10.1016/j.aca.2021.339342. Epub 2021 Dec 3. PMID: 35569842.
5. Qian Sun, Kerry J Welsh, David E Bruns, David B Sacks, Zhen Zhao, Inadequate Reporting of Analytical Characteristics of Biomarkers Used in Clinical Research: A Threat to Interpretation and Replication of Study Findings, Clinical Chemistry, Volume 65, Issue 12, 1 December 2019, Pages 1554–1562,
6. Clinical and Laboratory Standards Institute. (2012). Protocols for determination of limits of detection and limits of quantitation. CLSI-EP-17-A. Clinical and Laboratory Standards Institute.
7. Winek CL, Wahba WW, Windisch RM, Winek CL Jr. Serum alcohol concentrations in trauma patients determined by immunoassay versus gas chromatography. Forensic Sci Int. 2004;139(1):1-3.

Research Article

Outcomes of Sweat Conductivity Testing and Referral Patterns for Cystic Fibrosis: A 10-Year Retrospective Single-Center Study in Albania

Hamide Shllaku-Sefa[1*], Teuta Dedej-Kurti[2], Ndok Marku[1]

[1]*Laboratory Department, Catholic Hospital "Our Lady of Good Counsel", Tirana, Albania*
[2]*Laboratory Department, Faculty of Medicine, University of Medicine, Tirana, Albania*

Article Info

[*]***Corresponding Author:***
Hamide Shllaku-Sefa
Laboratory Department, Catholic Hospital "Our Lady of Good Counsel"
Street "Dritan Hoxha", Tirana, Albania
E-mail: sefadea@gmail.com
Phone.: 00355695305242

Keywords

Albania, Cystic Fibrosis, Immunoreactive Trypsinogen (IRT), Quantity Not Sufficient (QNS), Sweat Chloride Test

Abstract

Introduction: Cystic fibrosis (CF) is an autosomal recessive disorder. Sweat testing is a cornerstone of CF diagnosis and remains essential in settings where nationwide screening and registries are limited.

Materials and Methods: We performed a retrospective single-center analysis of all patients referred for sweat testing in our laboratory between May 2015 and May 2025 (n=567). Sweat induction and collection were performed using the Macroduct® system, and analysis was conducted via the Sweat-Chek™ Analyzer (Model 3120, ELITechGroup). Statistical analysis was performed using Jamovi software (version 2.3.28).

Results: The cohort included 302 males (53%) and 265 females (47%), with a mean age of 36.3 ±49.2 months (Median: 14; IQR: 4–48). The primary indications for referral were respiratory manifestations (57%), followed by gastrointestinal symptoms (14.1%), failure to thrive (6.2%), and others (4.2%). Asymptomatic individuals with no symptoms recorded at referral (18.2%). Diagnostic outcomes revealed that 83.4% of participants had normal results, 11.3% were positive, and 3% were borderline. A "quantity not sufficient" (QNS) rate of 2.3% was observed.

Conclusions: In this single-center referral cohort, most sweat test results were normal, with an overall positive rate of 11.3% and low QNS rates that are benchmark favorably against published quality targets. Clear referral pathways and strengthened early-life evaluation (including expansion of screening access) may support earlier diagnosis and reduce disease burden.

Introduction

Cystic fibrosis (CF), an autosomal recessive disorder caused by biallelic pathogenic variants in the CFTR gene, disrupts epithelial ion transport, leading to viscous mucus accumulation, chronic multi-organ inflammation, and progressive respiratory failure [1]. The dysfunctional CFTR protein impairs chloride/bicarbonate secretion and sodium hyperabsorption, resulting in hallmark manifestations: recurrent pulmonary infections, pancreatic insufficiency, elevated sweat chloride concentrations, and male infertility due to obstructive azoospermia [2].

Globally, CF incidence varies significantly by population. While Ireland reports 1:1,353 live births [3], Finland observes 1:25,000 [4], with Western Europe averaging 1:4,500 [5] and Northern/Central Europe 1:6,000 [6]. In Albania, a Southeastern European nation with limited epidemiological data, CF prevalence was historically underreported due to fragmented registries. Before 2017, no national CF registry existed; however, the Department of Pediatrics at the University Hospital Center "Mother Teresa" (QSUNT) served as the sole national referral center, systematically documenting all pediatric CF cases since 1992. Retrospective analysis of QSUNT records (1992–2017) estimated an incidence of 1:4,041 live births [7], though this figure reflects hospital-based ascertainment rather than population-wide screening and requires validation through prospective studies. Importantly, the incidence figure is an estimate derived from QSUNT historical case counts and national birth rates, and it is not calculated from the referral cohort in the present study (which includes patients referred for evaluation rather than a population-based denominator).

The sweat chloride test (SCT) remains the diagnostic gold standard per Cystic Fibrosis Foundation guidelines [8], quantifying electrolyte concentrations following pilocarpine iontophoresis. Despite technical standardization, preanalytical errors, particularly insufficient sample volume (QNS), compromise reliability, with reported QNS rates ranging from 5–25% globally [9,10]. In resource-limited settings, suboptimal patient preparation (e.g., dehydration, skin barriers) and inconsistent collection protocols exacerbate these challenges. This study addresses a critical gap: no prior research has characterized SCT outcomes, referral patterns, or QNS determinants in Albania's high-prevalence cohort. Over a decade (2015–2025), we analyzed 567 SCTs performed at our laboratory to:

Quantify sweat chloride distributions across age groups and clinical indications,

Evaluate temporal trends in referral practices following updated international diagnostic criteria,

Identify modifiable preanalytical factors contributing to QNS events.

Our findings provide actionable insights to optimize diagnostic pathways, reduce diagnostic delays, and support evidence-based implementation of newborn screening in similar underserved regions.

Materials and Methods

Study design and participants

This was a retrospective single-center study including all consecutive patients referred for sweat testing in our laboratory between May 2015 and May 2025. Inclusion criteria were: (i) referral for sweat testing during the study period and (ii) availability of key variables (age, sex, referral indication, and sweat test outcome). Exclusion criteria were: (i) duplicate records/repeat tests for the same individual during the study period (only the last test was included in the analysis) and (ii) missing key variables required for analysis."

Definition of referral indication categories

Referral indications were classified based on referral documentation and clinical notes. "Healthy individuals" refers to patients with no symptoms recorded at referral who were nevertheless referred for CF evaluation (e.g., screening or clinician concern without documented symptoms).

Definition of age categories

Although the study included patients aged 1 month to 39 years, the vast majority were pediatric patients. Subgroup analysis was performed based on age categories (≤3 months and >3 months), but no formal comparison between pediatric and adult populations was conducted due to the very small number of adult subjects.

We used the SS-032 Macroduct Supply Kit (Pilogel Discs, Macroduct Collectors), Macroduct Sweat Collection System model 3700-SYS, and Sweat-Chek 3120 Sweat Conductivity Analyzer by Eli Tech Group Biomedical Systems. The test was conducted in the following phases [11]:

Phase one – stimulation of sweating

The flexor surface of the child's forearm is cleaned with alcohol and deionized water and dried. Pilogel iontophoretic discs are applied to the skin with two stainless steel electrodes on the top, then fixed to the forearm. The electrodes are connected to a power supply, which automatically delivers pilocarpine equivalent to five minutes of iontophoresis at 1.5 mA, in accordance with international guidelines for sweat testing in both infants and adults [10].

Phase two – collection of sweat

After sweat stimulation, the pilogel discs are removed by hand, and the skin is cleaned with water. Next, the disposable collector for sweat collection is fixed to the stimulated area, and sweat is collected for 30 minutes, during which the child can move, play, eat, and drink. When the sweat collection is over, the collector is removed from the child's forehand. The sweat specimen is handled with care before analysis to

avoid introducing air bubbles into the sweat column.

Phase three – analytical phase

The chloride concentration in the collected sweat is measured in the laboratory. The end of the specimen tube is connected to the inlet nipple of the conductivity cell in the left-hand nipple. The sweat specimen is transferred from the collector tube into the take-up tube via the conductivity cell. After the sweat specimen is transferred into the cell, its electrical conductivity is measured, the electrolyte concentration is calculated, and the result appears on the digital display.

Statistics

Categorical variables are presented as frequency and percentage. Continuous variables are summarized as median (range). Group comparisons for categorical variables were performed using Fisher's exact test when expected cell counts were <5; otherwise, chi-square tests were used. A p-value ≤0.05 was considered statistically significant. Analyses were performed using Jamovi Statistical Software version 2.3.28.

Results

A total of 567 patients were enrolled in our study. There was a slight predominance of males with 53%. The mean age of the participants was 36.3 ± 49.2 months. Given the non-normal distribution of the data (Shapiro-Wilk p<.001), the median age of 14 months (IQR: 4–48 months) provides a more accurate representation of the cohort's central tendency.

The study analyzed a cohort primarily composed of patients older than 3 months (80%) and outpatients (69%). Clinical referrals were driven primarily by the Department of Pediatric Pulmonology (57%), with respiratory tract manifestations as the most common clinical sign for testing (57.3%). Detailed demographic and referral distributions are provided in Table 1.

Table 1: Descriptive statistics of the participants.

	N	%
Gender		
Males	302	53
Females	265	47
Age		
≤ 3 months	116	20
> 3 months	451	80
QNS		
No	554	97.7
Yes	13	2.3
Result		
Normal	473	83.4
Positive	64	11.3
Borderline	17	3
No results	13	2.3
Department		
Pediatric Pulmonology	321	57
General Pediatric Policlinic	90	16
Pediatric Gastroenterology	66	12
Pediatric Immunology-Allergy	42	7
Pediatric Intensive Care	22	4
Neonatal Department	7	1
Others	19	3
Signs and Symptoms		
Respiratory tract manifestations	325	57.3
Gastrointestinal tract manifestations	80	14.1

	N	%
Healthy individuals	70	12.3
Failure to thrive	35	6.2
Family history	13	2.3
IRT +	11	2
Confirmed mutations in the CFTR gene	9	1.6
Other manifestations	24	4.2
Hospitalization		
Inpatients	177	31
Outpatients	390	69

While the majority of tests yielded normal results (83.4%), a definitive positive diagnosis was reached in 11.3% of the cases. Statistical analysis revealed that a positive cystic fibrosis diagnosis was significantly associated with age, clinical symptoms, and hospitalization status ($p < .001$), whereas gender is not a relevant factor ($p = 0.346$). Table 2 provides the full breakdown of these diagnostic correlations.

Table 2: Demographic and clinical characteristics of participants by diagnostic result.

	Result								p
	Overall		Positive		Normal		Borderline		
	N	%	N	%	N	%	N	%	
Age									
≤ 3 months	112	20	26	23.2	85	75.9	1	0.9	<.001
> 3 months	442	80	38	8.6	388	87.8	16	3.6	
Gender									
M	294	53	39	13.3	245	83.3	10	3.4	0.377
F	260	47	25	9.6	228	87.7	7	2.7	
Signs and Symptoms									
Respiratory tract manifestations	318	57.4	25	7.9	283	89	10	3.1	<.001
Gastrointestinal tract manifestations	78	14.1	10	12.8	65	83.3	3	3.9	
Healthy individuals	70	12.6	2	2.9	67	95.7	1	1.4	
Failure to thrive	33	6	3	9.1	30	90.9	0	0	
Family history	13	2.3	3	23.1	10	76.9	0	0	
IRT +	11	2	4	36.4	6	54.5	1	9.1	
Confirmed mutations in the CFTR gene	9	1.6	8	88.9	0	0	1	11.1	
Other manifestations	22	4	9	41	12	54.5	1	4.5	
Hospitalization									
Inpatients	169	30	32	19	130	77	7	4	<.001
Outpatients	385	70	32	8.3	343	89.1	10	2.6	

Notably, in the subgroup of patients with confirmed CFTR mutations, 88.9% returned positive sweat tests, with the remaining 11.1% showing borderline results due to active modulator therapy (Table 2).

Among infants referred following a positive IRT screening, only 36.4% were confirmed positive by sweat testing, while over 54% returned normal results (Table 2). This highlights the critical role of SCT in resolving false positives from newborn screening.

Table 3 compares clinical characteristics by hospitalization status for patients with positive sweat test results. Analysis shows that signs and symptoms (p=0.005) represent the only variable with a statistically significant difference between inpatients and outpatients.

Table 3: Comparison of age, gender, and symptoms between inpatients and outpatients with positive sweat tests.

	Hospitalization						p
	Overall		Inpatients		Outpatients		
	N	%	N	%	N	%	
Age							
≤ 3 months	26	41	13	50	13	50	1.000
> 3 months	38	59	19	50	19	50	
Gender							
M	39	61	17	44	22	56	0.305
F	25	39	15	60	10	40	
Signs and Symptoms							
Respiratory tract manifestations	25	39	12	48	13	52	0.005
Gastrointestinal tract manifestations	10	16	9	90	1	10	
Healthy individuals	2	3	0	0	2	100	
Failure to thrive	3	5	2	67	1	33	
Family history	3	5	0	0	3	100	
IRT +	4	6	1	25	3	75	
Confirmed mutations in the CFTR gene	8	12	1	12.5	7	87.5	
Other manifestations	9	14	7	78	2	22	

The overall rate of insufficient sweat volume was low at 2.3% (n=13). Analysis of sample failure showed:

Non-significant factors: Age (p=0.351), gender (p=0.545), and specific clinical symptoms (p=0.315) did not significantly affect the ability to collect sufficient sweat.

Significant factors: A statistically significant difference was seen based on patient location (p=0.016), with inpatients exhibiting a higher QNS rate (4.5%) compared to outpatients (1.3%).

Detailed QNS distributions across all variables are summarized in Table 4.

Table 4: The distribution of the patients according to sweat volume, age, gender, symptoms and hospitalization.

	Sweat volume						p
	Overall		QNS		Sufficient		
	N	%	N	%	N	%	
Age							
≤ 3 months	116	20	4	3.4	112	96.6	0.315
> 3 months	451	80	9	2	442	98	
Gender							
M	302	53	8	2.6	294	97.4	0.587
F	265	47	5	2	260	98	

	Sweat volume						p
	Overall		QNS		Sufficient		
	N	%	N	%	N	%	
Signs and Symptoms							
Respiratory tract manifestations	325	57.3	7	2.2	318	97.8	0.315
Gastrointestinal tract manifestations	80	14.1	2	2.5	78	97.5	
Healthy individuals	70	12.3	0	0	70	100	
Failure to thrive	35	6.2	2	5.7	33	94.3	
Family history	13	2.3	0	0	13	100	
IRT +	11	2	0	0	11	100	
Confirmed mutations in the CFTR gene	9	1.6	0	0	9	100	
Other manifestations	24	4.2	2	8.3	22	91.7	
Hospitalization							
Inpatients	177	31	8	4.5	169	95.5	0.029
Outpatients	390	69	5	1.3	385	98.7	

Discussion

Cystic fibrosis is an autosomal recessive disease, common in Europe, including the Albanian population, but a national cystic fibrosis registry was not established in Albania until 2017 [7].

Sweat testing services have been provided since 2015, with an annual average of 57 tests. To ensure competency and quality, all tests are performed by a single laboratory professional. This volume exceeds the minimum requirements set by the Cystic Fibrosis Foundation and other professional bodies, which recommend at least 50 tests per laboratory and 10 per technician annually. [12].

While global literature often indicates a more severe disease trajectory in females, often attributed to the onset of estrogen-related complications post-puberty [13], our data did not reflect significant gender-based disparities in clinical presentation. We saw a slight male predominance (53%) and a higher positivity rate in males (13.3%) compared to females (9.6%), yet no significant differences were found in symptoms ($p=0.879$) or hospitalization rates ($p=0.204$). This lack of divergence is consistent with a prepubescent cohort, where the hormonal drivers typically associated with gender-specific CF morbidity have not yet manifested. Furthermore, since our study used cross-sectional data at the point of diagnosis without longitudinal follow-up, these findings represent the initial clinical state rather than long-term disease progression. Consequently, our results suggest that in the early pediatric stages of CF in Albania, gender does not appear to be a primary determinant of acute clinical severity.

The diagnoses of CF can be made at various ages, but with newborn screening, it's often diagnosed within the first month of life, before symptoms appear [14]. Most children are diagnosed by age 2, either through newborn screening or when symptoms become clear [15]. However, some individuals, particularly those with milder or atypical forms of the disease, may not be diagnosed until adolescence or adulthood [16]. Although we do not have a national program for newborn screening (NBS), since this is only offered in private hospitals in Albania, we have found a significant difference between the diagnosis of cystic fibrosis and age ($p<.001$) where it resulted that 23.2% of patients with a positive diagnosis were ≤3 months and only 8.6% were >3 months. The data prove that diagnosis typically occurs within the first three months of life, despite the lack of a universal screening program. This early detection points to effective diagnostic pathways and strong awareness among both caregivers and medical professionals. Most referrals originated from the Pediatric Pulmonology Department (57%), followed by General Pediatrics (16%), Gastroenterology (12%), and Immunology-Allergy (7%), with the remainder from Intensive Care (4%) and Neonatology (1%). These patterns diverge from international guidelines, such as those of the Croatian Society of Medical Biochemistry and Laboratory Medicine and University Hospital Centre Zagreb, which advocate for primary care physicians to initiate referrals to CF specialists [17]. Furthermore, our findings contrast with data from Turkey, where general pediatric clinics serve as the primary source of referrals [18]. We suggest these discrepancies reflect distinct national variations in diagnostic and management strategies for cystic fibrosis.

Since 1992, patients with CF in Albania have been managed and followed at the Department of Pediatrics within the University Hospital Center "Mother Teresa". The CF diagnosis

is based predominantly on clinical criteria and two positive sweat test results for patients suspected of the disease. The diagnosis is confirmed by the identification of two CF-causing mutations. Genetic assessment has been in practice since 2004, when the search for genetic mutations in patients with cystic fibrosis began to be recorded. Unfortunately, there was no national registry for cystic fibrosis in Albania until 2017. However, every pediatric patient diagnosed in the Department of Pediatrics at QSUNT has been documented throughout the course of the disease. Practically, this is the only center that diagnoses and follows cystic fibrosis patients in our country. Since 2017, Albania has been part of the European Cystic Fibrosis Society Patient Registry (ECFSPR), which collects anonymized demographic and clinical data from consenting people in Europe with CF [7].

Our study confirms that respiratory and gastrointestinal manifestations are still the primary drivers for sweat test referrals, accounting for 57.4% and 14.1% of our cohort, respectively. The strong correlation between age, symptoms, and test results ($p<.001$) aligns with established literature [2] about the early clinical presentation of cystic fibrosis.

Beyond clinical symptoms, other primary indications for sweat test referral by pediatricians include family history, positive IRT results on newborn screening, and confirmed CFTR mutations via molecular analysis.

In our cohort, only 23.1% of patients with a family history of CF yielded positive sweat test results. While family history indicates an increased carrier probability, and thus a higher risk of having an affected child, the actual clinical risk depends on the specific CFTR mutations present in the family. Consequently, genetic counseling and molecular testing are essential to determine the precise reproductive risk for individuals and their offspring [19].

Approximately 97% of infants with a single CFTR mutation identified through newborn screening for immunoreactive trypsinogen are false positives [20]. Consequently, CF has become a primary model in the United States for integrating DNA analysis into NBS programs [21]. While genetic-based NBS can find presymptomatic infants, it also finds asymptomatic heterozygote carriers and infants without disease-causing mutations. In such cases, at least one parent is typically a carrier, placing the couple at risk for future affected pregnancies [22]. While false-positive rates vary by screening methodology and ethnic background [23], our data revealed a false-positive IRT rate of 54.5%. It is critical to clarify that these referrals do not stem from a centralized national screening program, as a universal NBS for cystic fibrosis is not yet implemented in the Albanian public health sector. Instead, these cases represent a specific population of infants born in private maternity hospitals where elective screening is available. This distinction is vital for interpreting our data, as it reflects a self-selected or higher-socioeconomic demographic rather than a cross-section of the general population. Our findings underscore that even in these private settings, the SCT remains the indispensable "gold standard" in resolving the diagnostic uncertainty created by initial IRT elevations.

While it may initially appear that identifying CFTR mutations is sufficient for diagnosis, the sweat chloride test remains the gold standard for confirming a clinical diagnosis by verifying the functional defect of the CFTR protein [24,25]. This is particularly critical in cases involving a single identified mutation or indeterminate results [26]. In our study, nine patients presented with confirmed F508del mutations; eight yielded positive sweat test results, while one returned a borderline result. The latter case involved a patient previously diagnosed with CF who had been receiving CFTR modulator therapy in the United Kingdom for nearly a year. This case underscores an evolving role for the SCT, moving beyond initial diagnosis to become a key metric for assessing the effectiveness of CFTR modulators, which work by enhancing salt and water transport across cell membranes [27].

Children with cystic fibrosis experience significantly higher rates of hospitalization and outpatient visits compared to children without CF, with inpatient hospitalization rates often exceeding 30 times those of their peers [28]. In our cohort, patients with a positive SCT had a hospitalization rate of 19%, which was more than double the outpatient management rate of 8.3%. Additionally, we found a statistically significant association between positive SCT results and hospitalization ($p < 0.001$). These findings call for more research on diagnosis, care, and treatment to improve patient outcomes, prevent complications, and reduce hospitalizations.

The determination of QNS rates serves as a primary metric for assessing the quality of pre-analytical sweat collection. The QNS rates observed in our laboratory were well below the thresholds recommended by the Cystic Fibrosis Foundation (<10% for infants ≤3 months and <5% for older children) [10], reflecting strong adherence to preanalytical quality standards. These results provide strong evidence of the high technical proficiency of our laboratory staff. Furthermore, while literature suggests that physiological variables like body weight or skin condition can influence volume [2, 18, 29], our statistical analysis demonstrated that age ($p=0.315$), gender ($p=0.587$), and clinical symptoms ($p=0.315$) did not compromise sample adequacy. This stability across diverse pediatric subgroups underscores the robustness of our standardized collection protocols.

Conclusions

In this single-center referral cohort, sweat testing yielded normal results, with an overall positive rate of 11.3% and low QNS rates that benchmark favorably against published targets. Expanding access to early-life evaluation (including broader screening availability), strengthening referral pathways, and

maintaining high-quality sweat testing procedures may support earlier CF diagnosis and improved outcomes.

Disclosure

The authors have nothing to declare.

Ethical Approval

Our study involved human subjects and followed the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Authors contributions

Hamide Shllaku-Sefa contributed to conceptualization, data curation, formal analysis, investigation, validation, methodology, visualization, and writing – original draft. Ndok Marku, role in conceptualization, data curation, supervision, and project administration. Teuta Dedej-Kurti contributed to writing - review & editing.

Funding sources

None.

Data Availability Statement

Data will be provided on request.

References

1. Davis PB. Cystic fibrosis since 1938. Am J Respir Crit Care Med. 2006;173(5):475-482.
2. Farrell PM, Rosenstein BJ, White TB, Accurso FJ, Castellani C, Cutting GR, et al. Guidelines for diagnosis of cystic fibrosis in newborns through older adults: cystic Fibrosis Foundation consensus report. J Pediatr 2008;153: S4–S14.
3. Farrell PM, Joffe S, Foley L, Canny GJ, Mayne P, Rosenberg M. Diagnosis of cystic fibrosis in the Republic of Ireland: Epidemiology and costs. Ir. Med. J. 2007; 100:557–560.
4. Kere J, Estivill X, Chillon M, Morral N, Nunes V, Norio R, et al. Cystic fibrosis in a low-incidence population: Two major mutations in Finland. Hum. Genet. 1994; 93:162–166.
5. Dankert-Roelse JE, Bouva MJ, Jakobs BS, Janssens HM, de Winter-de Groot KM, Schonbeck Y, et al. Newborn blood spot screening for cystic fibrosis with a four-step screening strategy in the Netherlands. J. Cyst. Fibros. 2019; 18:54–63.
6. Skov M, Baekvad-Hansen M, Hougaard DM, Skogstrand K, Lund AM, Pressler T, et al. Cystic fibrosis newborn screening in Denmark: Experience from the first 2 years. Pediatr. Pulmonol. 2020; 55:549–555.
7. Kasmi I, Kasmi G, Basholli B, Sefa HS, Vevecka E. The Spectrum and Frequency of Cystic Fibrosis Mutations in Albanian Patients. Balkan J Med Genet. 2024;27(1):31-36.
8. Castellani C, Southern KW, Brownlee K, Dankert Roelse J, Duff A, Farrell M, et al. European best practice guidelines for cystic fibrosis neonatal screening. J Cyst Fibros 2009;8(3):153–173.
9. Coakley J, Scott S, Doery J, Greaves R, Talsma P, Whitham E, et al. Australian guidelines for the performance of the sweat test for the diagnosis of cystic fibrosis: report from the AACB Sweat Testing Working Party. Clin Biochem Rev 2006;27(2): S1–7.
10. LeGrys VA, Yankaskas JR, Quittell LM, Marshall BC, Mogayzel PJ Jr. Cystic Fibrosis Foundation. Diagnostic sweat testing: the Cystic Fibrosis Foundation guidelines. J Pediatr 2007;151(1):85–89.
11. ELITechGroup. 3120 SWEAT CHEK™ Sweat Conductivity Analyzer Instruction Manual. Logan, UT: Wescor, Inc.; 2005.
12. Massie J, Greaves R, Metz M, Wiley V, Graham P, Shepherd S, et al. Australasian Guideline (2nd Edition): an Annex to the CLSI and UK Guidelines for the Performance of the Sweat Test for the Diagnosis of Cystic Fibrosis. Clin Biochem Rev. 2017;38(3):115-130.
13. McNally E, Casey M. Sex disparities in cystic fibrosis in the era of highly effective modulator treatment. BMC Pulm Med. 2025;25(1):212.
14. Borowitz D, Robinson KA, Rosenfeld M, Davis SD, Sabadosa KA, Spear SL, et al. Cystic Fibrosis Foundation evidence-based guidelines for management of infants with cystic fibrosis. J Pediatr. 2009;155(6 Suppl): S73-93.
15. A.D.A.M. Medical Encyclopedia [Internet]. Cystic Fibrosis; [reviewed 2024 Feb 17; cited 2025 Sep 03]; [about 5 p.]. Available from: https://medlineplus.gov/ency/article/000107.htm#:~:text=Many%20people%20carry%20a%20CF,Symptoms%20in%20newborns%20may%20include: (accessed:30/10/2025)
16. Padoan R, Quattrucci S, Amato A, Carnovale V, Salvatore D, Salvatore M et al. The Diagnosis of Cystic Fibrosis in Adult Age. Data from the Italian Registry. Diagnostics (Basel). 2021;11(2):321.
17. Leničeк Krleža J, Aralica M, Tješić-Drinković D, Crneković K, Culej J, Fressl Juroš G, et al. National Guidelines for the Performance of the Sweat Test in Diagnosis of Cystic Fibrosis on behalf of the Croatian Society of Medical Biochemistry and Laboratory Medicine and the Cystic Fibrosis Centre - Paediatrics and adults, University Hospital Centre Zagreb. Biochem Med. 2022;32(1):010501.
18. Ozgenc A, Yasar NE, Kucur M, Bolayirli I M, Konukoglu D. The most common preanalytic problem of sweat testing: Insufficient sweat volume. Int J Med Biochem. 2020; 3(2), 62-66.
19. Mishra A, Greaves R, Massie J. The Relevance of Sweat testing for the Diagnosis of Cystic Fibrosis in the Genomic Era. Clin. Biochem. Rev. 2005; 26, 135–153.
20. Rock MJ, Hoffman G, Laessig RH, Kopish GJ, Litsheim TJ, Farrell PM. Newborn screening for cystic fibrosis in

Wisconsin: Nine-year experience with routine trypsinogen/DNA testing. Journal of Pediatrics. 2005;147(Suppl. 3): S73–S77.
21. Sharp JK, Rock MJ. Newborn screening for cystic fibrosis. Clinical Reviews in Allergy and Immunology. 2008; 35:107–115.
22. Tluczek A, Orland KM, Cavanagh L. Psychosocial consequences of false-positive newborn screens for cystic fibrosis. Qual Health Res. 2011;21(2):174-186.
23. Giusti R. Elevated IRT levels in African American infants: Implications for newborn screening in an ethnically diverse population. Pediatric Pulmonology. 2008; 43:638–641.
24. Farrell PM, White TB, Ren CL, Hempstead SE, Accurso F, Derichs N, et al. Diagnosis of Cystic Fibrosis: Consensus Guidelines from the Cystic Fibrosis Foundation. J Pediatr. 2017;181 S: S4-S15: e1.
25. Massie J, Greaves R, Metz M, Wiley V, Graham P, Shepherd S, et al. Australasian Guideline (2nd Edition): an Annex to the CLSI and UK Guidelines for the Performance of the Sweat Test for the Diagnosis of Cystic Fibrosis. Clin Biochem Rev. 2017;38(3):115-130.
26. Mishra A, Greaves R, Massie J. The relevance of sweat testing for the diagnosis of cystic fibrosis in the genomic era. Clin Biochem Rev. 2005;26(4):135-153.
27. Zemanick ET, Ramsey B, Sands D, McKone EF, Fajac I, Taylor-Cousar JL, et al. Sweat chloride reflects CFTR function and correlates with clinical outcomes following CFTR modulator treatment. J Cyst Fibros. 2025;24(2):246-254.
28. Malmivaara K, Pölkki M, Prami T, Raittinen P, Heikkilä E, Aalto A, et al. Cystic fibrosis patient characteristics and healthcare resource utilization in Finland using linked registries. Heliyon. 2024. 22;10(13): e33439.
29. Camargos P, Bedran RM, Alvim CG, Sader OG. Rate and predictors of insufficient sweat volume in very young infants after pilocarpine gel iontophoresis: prospective, population-based study. Arch Dis Child. 2024;109(4):304-307.

Research Article

Reassessing Vitamin D Deficiency: Data Driven Indirect Reference Intervals for 25-Hydroxyvitamin D in Pakistan via refineR Algorithm

Sibtain Ahmed[1*], Amal Mahmood[2], Ashishkumar Agravatt[3], Imran Siddiqui[1], Aysha Habib Khan[1]

[1]*Section of Chemical Pathology, Department of Pathology & Laboratory Medicine, the Aga Khan University, Karachi, Pakistan*
[2]*Medical College, the Aga Khan University, Karachi, Pakistan*
[3]*PDU Medical College, Rajkot, India*

Article Info

[*]***Corresponding Author:***
Sibtain Ahmed
Assistant Professor, Section of Chemical Pathology
Department of Pathology and Laboratory Medicine
Aga Khan University, Karachi, Pakistan
E-mail: Sibtain.ahmed@aku.edu

Keywords

Vitamin D, Reference interval, Pakistan Classification Description, Metabolism, Endocrinology, Vitamin D

Abstract

Objectives: To establish population-specific reference intervals (RIs) for serum 25-hydroxyvitamin D [25(OH) D] in the Pakistani population using an indirect big data approach, and to evaluate the prevalence of vitamin D deficiency across age and sex strata.

Methods: This retrospective study analyzed serum 25(OH) D levels from 201,625 individuals aged ≥18 years who underwent testing at a tertiary care center over five years. Data were extracted from electronic medical records. Serum 25(OH)D was measured via chemiluminescent immunoassay (DiaSorin Liaison), with internal and external quality controls. The refineR algorithm with modBoxCox transformation was used to derive RIs. Statistical analysis was conducted using R software.

Results: The median serum 25(OH)D level was 17.2 ng/mL, with 25.7% of individuals classified as deficient (≤20 ng/mL), and 33% severely deficient (≤12 ng/mL). Vitamin D deficiency was more pronounced in men and younger adults. Reference intervals derived using modBoxCox transformation were 10.3–51.5 ng/mL for females and 6.1–50.6 ng/mL for males. Notably, females demonstrated higher vitamin D levels across all age groups compared to males.

Conclusion: This study presents large-scale, data-driven estimated reference intervals for 25(OH)D in Pakistan using an indirect approach. The findings highlight a high burden of vitamin D deficiency, particularly among men and younger adults, and reinforce the need for localized reference values to improve diagnostic accuracy. These insights support the development of targeted screening and supplementation strategies, and the reevaluation of clinical thresholds based on population specific data.

Introduction

Vitamin D is fat soluble vitamin that essential for maintaining physiological homeostasis, particularly in calcium and phosphorus metabolism, and bone health through the regulation of bone mineralization. Beyond its classical role in skeletal function, vitamin D exerts significant immunomodulatory effects by influencing the activity of T and B lymphocytes. These actions suggest broader function of vitamin D in managing both innate and adaptive immune mechanisms. A growing body of evidence suggests that insufficient levels of vitamin D may contribute to the development of several autoimmune conditions such as multiple sclerosis and type 1 diabetes by affecting immune regulation and promoting inflammatory responses [1].

Vitamin D is primarily found in two variants: ergocalciferol (D2) and cholecalciferol (D3). The D2 form is made when ultraviolet B (UVB) radiation interacts with ergosterol, a compound present in fungi, yeast, and certain plants. It can also be obtained from plant-derived sources such as mushrooms. In contrast, vitamin D3 is formed in the skin when 7-dehydrocholesterol is exposed to UVB radiation, and it is also present in animal-derived foods such as cod liver oil. After entering the body, whether obtained from the diet or produced by the skin, vitamin D2 and D3 (collectively termed calciol) attach to a transport protein known as vitamin D-binding protein (VDBP) in the bloodstream. This complex then travels to the liver, where vitamin D undergoes its first hydroxylation step, converting it into 25-hydroxyvitamin D, the primary circulating form found in blood. Subsequently, a second hydroxylation occurs in the kidneys, resulting in the formation of calcitriol, or 1,25-dihydroxyvitamin D the biologically active version which then binds to vitamin D receptors to carry out its functions [2].

Different health authorities have issued varying recommendations for vitamin D, but a common consensus is that most people maintain adequate levels when their serum 25-hydroxyvitamin D [25(OH)D] concentration reaches at least 50 nmol/L (equivalent to 20 ng/mL). According to the National Academies of Sciences, Engineering, and Medicine, vitamin D levels under 12 ng/mL (30 nmol/L) are classified as deficient, while concentrations ranging from 12 to 20 ng/mL (30–50 nmol/L) may not be sufficient for certain individuals [3]. The European Food Safety Authority (EFSA) and the Nordic Nutrition Recommendations also support a threshold of 50 nmol/L as adequate for the general population [4]. The Endocrine Society previously categorized vitamin D deficiency in 2011 as having serum 25(OH)D levels below 20 ng/mL (50 nmol/L). However, in its 2024 revision, the organization shifted away from using a universal threshold, instead promoting a more personalized evaluation approach [5].

Research shows that low 25(OH)D are prevalent across the world. Approximately 40.4% of European individuals have serum 25(OH)D concentrations under 50 nmol/L, and roughly 13% experience more severe deficiency, with levels falling below 30 nmol/L (equivalent to 12 ng/mL) [6]. In the United States, around 5.9% of the population experiences severe 25(OH)D deficiency, while in Canada, the figure is approximately 7.4% [7,8]. In Pakistan, inadequate 25(OH) D levels are a major public health issue, influenced by factors such as insufficient sunlight exposure, cultural norms, and poor dietary intake across various segments of the population. Pakistan has the highest incidence of adult vitamin D deficiency in South Asia (73%), with a mean level of 17.93 ng/mL [9].

Clinicians depend on established reference intervals (RIs) to accurately interpret laboratory test results. Traditionally, these RIs have been determined using direct methods outlined by the Clinical and Laboratory Standards Institute (CLSI), which involve selecting a well-defined group of healthy individuals who meet strict criteria. The reference range is typically defined by the values between the 2.5th and 97.5th percentiles [10]. However, more recent advances favor indirect methods that analyze large datasets often exceeding 50,000 samples drawn from routine clinical data stored in laboratory information systems. These approaches have proven to be more precise than traditional methods that rely on just 120 samples. An innovative approach, known as the refineR algorithm, applies inverse modeling techniques to more precisely calculate reference intervals (RIs), even when data distributions are skewed or normal. Research indicates that refineR is capable of generating RIs that closely align with those obtained through traditional direct methods, making it a useful and efficient option for clinical laboratories [11].

There is a lack of population-specific reference intervals that accurately reflect the 25(OH)D status of healthy individuals in Pakistan. Most available reference values are based on direct sampling of limited, often non-representative populations, or extrapolated from Western data, which may not be appropriate for local demographics due to ethnic, environmental, and lifestyle differences. To address this gap, our study aims to leverage the power of big data and apply the RefineR algorithm, an indirect statistical approach, to establish more accurate and representative reference intervals for serum 25(OH)D in the Pakistani population. By analyzing a large dataset from routine laboratory results, this study seeks to establish evidence-based reference intervals that more accurately represent the overall health status of the general population.

Methods

Study design

This retrospective, observational study examined serum samples from individuals who had received inpatient or outpatient care at a tertiary hospital over the preceding five-year period from 01/01/2018 to 31/12/2022. This study was exempted by the institutional ethical committee of Aga Khan University (2023-8678-25064) and the study was performed using anonymized data set in accordance with the Declaration of Helsinki.

Data collection

Information including patient sex, age, date of testing, and clinical department was retrieved from electronic health records and systematically categorized. Initially, data from 273,240 individuals over a five-year period were considered. To avoid duplicate entries, if multiple 25(OH)D measurements were recorded for a patient on the same day, only the earliest result was retained, and subsequent ones were excluded. Before conducting statistical analysis, the dataset was refined by removing patients aged <18 and any entries lacking key demographic or clinical information. After removal of duplicates and data of patients less than 18, the remaining data included in the final analysis was 201,625 patients. As this was a retrospective study based on routine laboratory data, no explicit clinical exclusion criteria (e.g., chronic kidney disease, liver disease, malabsorption syndromes, or vitamin D supplementation) could be applied due to limitations in available metadata. Therefore, the dataset likely includes both healthy individuals and patients with underlying conditions affecting vitamin D status. However, the refineR indirect method is specifically designed to analyze mixed populations and estimate the underlying non-pathological distribution by minimizing the influence of extreme values.

Serum vitamin-D analysis

Serum 25(OH)D concentrations were measured using the Liaison analyzer (DiaSorin Inc., USA), which employs a chemiluminescent immunoassay method. The assay featured an analytical range of 4.0 to 150 ng/mL. Internal quality control was maintained by including two levels of control materials with each sample run. External quality assurance was ensured through participation in the College of American Pathologists (CAP) proficiency testing, conducted three times annually over the course of the study. All results from these external assessments were within acceptable performance criteria.

Data management and statistical analysis

Following data validation, statistical analyses were conducted using R software (version 4.4.1). The refineR package was utilized to estimate reference intervals (RIs) for serum 25(OH) D. Continuous data were described using either the mean and standard deviation (SD) or the median with interquartile range (IQR) depending on the distribution. Categorical variables were reported as counts and corresponding percentages.
The refineR algorithm accessible as a free R package is designed to derive RIs from real-world datasets that include both healthy and abnormal results. For datasets with skewed values, it uses the modified Box-Cox (modBoxCox) transformation, which incorporates an additional parameter, utilized to better normalize the data and improve RI accuracy [11].

Results

The mean age of the individuals in this study was approximately 55.5 years. Among the total population, females accounted for 68.51% (138,155 individuals), while males made up 31.49% (63,470 individuals) Table 1. The median serum level of 25(OH)D was measured at 17.2 ng/mL, with an interquartile range of 9.94 to 26.7 ng/mL. The mean concentration was calculated to be 21.58 ng/mL, with a standard deviation of ±17.28 ng/mL. According to the Endocrine Society's 2011 criteria, 25.7% (51,896 individuals) exhibited vitamin D deficiency (≤20 ng/mL or ≤50 nmol/L), 20% (40,393 individuals) were categorized as insufficient (21–29 ng/mL or 52.5–72.5 nmol/L), and 21.2% (42,725 individuals) had adequate levels (≥30 ng/mL or ≥75 nmol/L). Furthermore, a severe deficiency (≤12 ng/mL or ≤30 nmol/L) was observed in 33.03% (66,611 individuals) of the cohort (Table 2).

Table 1: Stratification of data according to age group and gender.

Age group	Female	Male	Total
18-44	76,619	27,939	104,558
45-65	47,627	25,820	73,448
>65	13,909	9,711	23,620
Total	**138,155**	**63,470**	**201,625**

Table 2: Percentage of serum [25(OH)D] results indicating severe deficiency, and insufficiency.

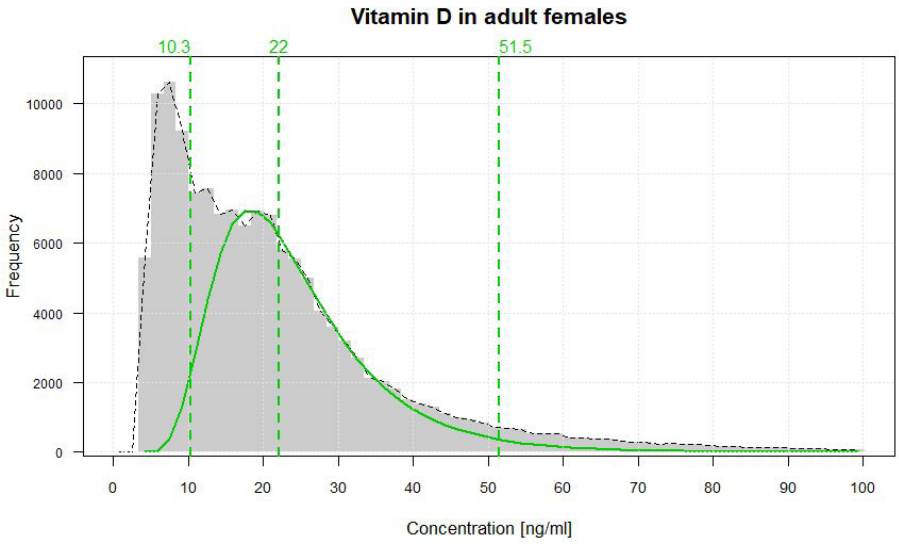
Age group	Vitamin D status	Females (76,619)	Males
18-44	Severe deficiency	29,959 (39.1%)	11,603(41.5%)
	Deficiency	18,319 (23.9%)	8,382 (30%)
	Insufficiency	13,409 (17.5%)	4,080 (14.6%)
	Sufficiencyw	14,932 (18.7%)	3,874 (13.6)
45-65	Severe deficiency	12,555 (26.4%)	7,937 (30.7%)
	Deficiency	11,800 (24.8%)	7,926 (30.69%)
	Insufficiency	11,547 (24.2%)	5,093 (19.7%)
	Sufficiency	11,725 (24.6%)	4,864 (18.9%)
>65	Severe deficiency	2,498 (18%)	2,059 (21.2%)
	Deficiency	2,911(20.9%)	2,558(26.3%)
	Insufficiency	3,763 (27.1%)	2,501 (25.8%)
	Sufficiency	4,737 (34%)	2,593 (26.7%)

A total of 201,625 serum 25(OH)D samples were analyzed using the refineR approach to determine optimal reference intervals (RIs). Due to the right-skewed nature of the dataset, standard statistical techniques were less appropriate. Therefore, a modified Box-Cox (modBoxCox) transformation was applied to our study cohort. The modBoxCox method adjusts for this imbalance by transforming the data into a shape that is closer to normal distribution, allowing for more precise calculation of reference intervals and better representation of the population's 25(OH)D status.

Females

For females (n = 138,155), the reference intervals derived from the overall dataset after applying the modBoxCox transformation showed a lower limit (LL) of 10.3 ng/mL (95% CI: 7.37–11.7 ng/mL) and an upper limit (UL) of 51.5 ng/mL (95% CI: 44.9–66.9 ng/mL). Distribution of 25(OH)D in females is depicted in Figure 1.

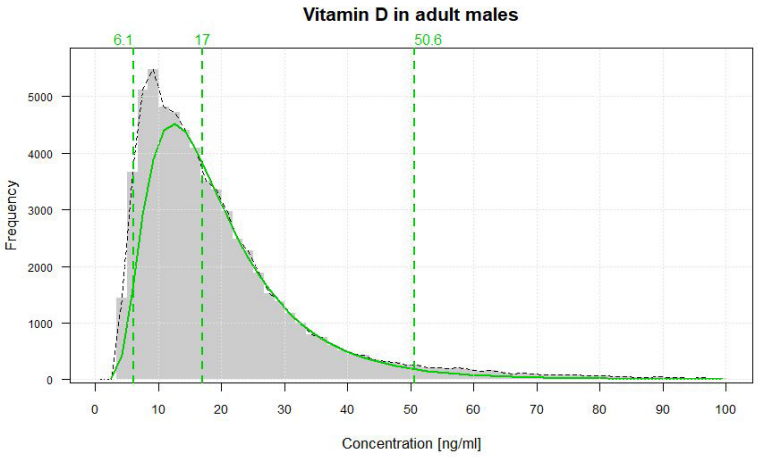
Table 1: Distribution of test results in females.



Males

For the male subgroup (n = 53,024), the estimated reference interval for serum 25(OH)D levels ranged from a lower limit of 6.1 ng/mL (95% CI: 2.36–7.36 ng/mL) to an upper limit of 50.6 ng/mL (95% CI 41.8–54.7ng/mL). Distribution of 25(OH) D in males is depicted in Figure 2.

Figure 2: Distribution of test results in Males.




```
Reference Intervals
-------------------------------------------------
lower limit [ 2.5% perc]: 6.1 (2.36; 7.36)
upper limit [97.5% perc]: 50.6 (41.8; 54.7)

Model Parameters
-------------------------------------------------
     method: refineR (v1.6.2)
      model: modBoxCox
     N data: 63470
N bootstrap: 200
    rounded: yes (base: 0.01)
  point est: fullDataEst
     lambda: 0.049
         mu: 2.92
      sigma: 0.696
      shift: 1.61
       cost: -22.9
NP fraction: 0.845
```

Discussion

Vitamin D deficiency is widely acknowledged as a major global health issue, affecting not only bone strength but also the immune system and the risk of developing long-term illnesses. In regions like South Asia, particularly Pakistan, deficiency rates remain alarmingly high across all age groups, despite abundant sunlight [9,12]. This paradox is often linked to several contributing factors, including insufficient exposure to sunlight, cultural norms involving modest clothing, inadequate dietary sources of vitamin D, and darker skin tones [13]. Although many studies have documented prevalence of vitamin D deficiency in Pakistan, there remains a clear gap in establishing standardized reference intervals tailored to the local population. Currently, most clinical evaluations are based on ranges derived from Western cohorts, which may not adequately account for the unique environmental and physiological characteristics of Pakistani population. This gap hampers accurate diagnosis, risk stratification, and the formulation of effective public health strategies.

This study presents a large-scale population-based analysis generating indirectly derived, data-driven reference interval estimates for 25(OH)D in Pakistan, utilizing data from over 200,000 individuals stratified by sex and age. The findings offer critical insights into the national burden of vitamin D deficiency, highlight notable sex- and age-based differences, and underscore the urgent need for targeted public health interventions.

When stratified by age, the data reveals that younger adults (18–44 years) constitute the largest proportion of the tested population. Among them, severe deficiency (<12 ng/mL) was alarmingly common, 39.1% of females and 41.5% of males. This is consistent with the available literature reporting that young adults have high prevalence of vitamin D deficiency [14]. These findings are concerning given that this age group represents the most economically productive segment of the population and suggests that early-life vitamin D supplementation or dietary interventions may be lacking.

In the 45–65 age group, the prevalence of severe deficiency slightly declined, possibly due to increased health awareness or supplementation with age. However, a significant proportion still remained either deficient or insufficient. The elderly group (>65 years) showed the lowest rates of severe deficiency, though this might reflect survivorship bias or greater medical attention and supplementation in this age group. Studies show that prevalence of vitamin D deficiency is less among the age group above 65 years [15,16]. Notably, sufficiency was most prevalent in this age group among females (34%). The vitamin D status reported in this Pakistani cohort is indicative of a widespread deficiency, with patterns similar to those observed in other South Asian countries [9]. However, the higher rates of severe deficiency among men is particularly striking and contrast with trends in other populations, where women typically show higher rates of deficiency [17]. These differences may stem from lifestyle, socioeconomic, and cultural factors unique to Pakistan, such as dietary patterns low in vitamin D-rich foods, limited sun exposure, and a lack of fortification policies.

To derive epidemiologically sound, population-based reference intervals (RIs) for serum 25(OH)D, we employed the refineR indirect method, modified Box-Cox (modBoxCox) transformations. This approach allows for modeling of non-pathological distributions within large clinical datasets by identifying the main distribution peak and estimating the

parameters of a transformed normal distribution. In our dataset, which included over 200,000 individuals, the distribution of vitamin D levels was right-skewed, warranting the application of a two-parameter modBoxCox transformation, in line with recommendations by Ammer et al. [18].

Using this approach, we obtained refined reference intervals for 25(OH)D across different age and sex strata. For the entire population, modBoxCox transformation yielded a lower reference limit (LRL) closer to the clinically significant deficiency threshold of 12 ng/mL, strengthening the clinical relevance of the calculated RIs. Specifically, as seen in our histogram-based plots (Figure 2), the modBoxCox-derived LRLs aligned with early signs of hypovitaminosis D, confirming its suitability for Pakistani populations where right-skewed distributions are common due to high deficiency prevalence. The calculated reference intervals demonstrate a clear sex-based disparity. Women (n = 138,155) exhibited higher lower limits (LL = 10.3 ng/mL) and upper limits (UL = 51.5 ng/mL) compared to men (n = 63,470), who showed an LL of 6.1 ng/mL and UL of 50.6 ng/mL. These findings are consistent with previous studies indicating gender-related differences in vitamin D metabolism [19]. The wider range and lower LRL in men suggest a greater prevalence of suboptimal vitamin D status among males, which is further supported by the distribution patterns observed in the histograms.

Notably, the reference intervals derived via modBoxCox transformation closely matched those obtained via standard transformation methods, but offered greater alignment with biological plausibility and clinical thresholds. For example, among females, the modBoxCox-derived LRL was 10.3 ng/mL, while for males it was 6.1 ng/mL, values that mirror the levels below which physiological symptoms may occur and are closer to the sever deficiency cut off (<12ng/ml). This reinforces the applicability and transferability of modBoxCox-derived RIs to clinical settings, especially in low-resource contexts where direct recruitment of healthy reference populations may not be feasible. A recent cross sectional study by Usama et al. (2024) focused on determining reference ranges for serum 25(OH) D levels tailored to the population of Peshawar, Pakistan. The researchers determined a reference range of 6.43–45.0 ng/mL (2.5th–97.5th percentile) for 25(OH)D. This study however utilized data from only 200 adults aged 18-54 from Peshawar, limiting the applicability of the data [20]. A study by Peralas-Afan et al. underscored the importance of seasonal differences in the vitamin D levels whereby the median vitamin d levels were higher in summer than winter. One of the reason ins possible low penetrance of UV light in winter season [21,22]. The authors suggest that 25(OH)D levels above 12.0 ng/mL may be sufficient for a healthy population, challenging the prevailing notion of a widespread vitamin D deficiency pandemic. This perspective advocates for a more refined interpretation of vitamin D sufficiency, potentially reducing unnecessary supplementation [22].

The findings have major implications for clinical practice and public health policy. First, the establishment of locally derived RIs provides a more accurate benchmark for diagnosing deficiency in Pakistani populations. Second, the high prevalence of deficiency and insufficiency in all age groups, particularly among men and younger adults, supports the need for routine screening, community education, and national supplementation programs. Moreover, these results call for a re-evaluation of current diagnostic thresholds used in clinical labs across Pakistan, which are often based on Western reference standards. A locally adapted approach to diagnosing and managing vitamin D deficiency could help mitigate long-term health outcomes, including bone disorders, immune dysfunction, and cardiometabolic diseases.

One of the key strengths of this research lies in its extensive sample size and the detailed analysis stratified by age and gender, enhancing the reliability and applicability of the results. The application of the modBoxCox transformation contributed to the accurate determination of reference intervals. Nonetheless, there are certain limitations to consider. This study has several limitations. First, the dataset was derived from a tertiary care center and likely includes symptomatic individuals and patients undergoing testing due to suspected vitamin D deficiency, introducing potential selection bias. As a result, the assumption of a strictly healthy reference population is not fully met. Second, due to the retrospective design and lack of detailed clinical metadata, we were unable to exclude individuals with conditions known to affect vitamin D levels, such as chronic kidney disease, liver disease, malabsorption syndromes, or use of vitamin D supplementation. These factors may have influenced the distribution of 25(OH)D concentrations. Although the refineR algorithm is designed to account for mixed populations by estimating the central non-pathological distribution, the derived intervals should be interpreted as statistical estimates rather than true physiological reference intervals. Lastly, seasonal variation in vitamin D levels was not accounted for and may represent an additional confounder. Since the study relied on retrospective data, there is a possibility of selection bias, particularly because individuals tested for vitamin D levels may have already been suspected of having deficiencies. Moreover, the influence of seasonal changes on vitamin D concentrations was not accounted for, indicating a potential area for investigation in future studies.

Conclusion

This study provides large-scale, data-driven estimated reference intervals for serum 25(OH)D in a Pakistani population using an indirect approach. While these findings offer important insights into population-level vitamin D status and highlight a high burden of deficiency, the intervals should be interpreted as indirectly derived statistical estimates rather than definitive physiological reference intervals. Further prospective studies with well-characterized healthy populations are required for external validation before clinical implementation

Ethical Consideration
This study was exempted by the institutional ethical committee of Aga Khan University (2023-8678-25064).

Consent
Not applicable as anonymized data was analyzed.

Conflicts of Interests
All authors declare no conflicts of interests.

Funding
None.

Data Availability
Anonymized data is available on reasonable requests from corresponding author.

Author contribution
Sibtain Ahmed contributed to study conceptualization and manuscript writing. Amal Mahmood contributed to the literature review and manuscript writing. Ashishkumar Agravatt performed data analysis and contributed to manuscript writing. Imran Siddiqui contributed to manuscript writing, table preparation, and figure development. Aysha Habib Khan contributed to manuscript writing and critical intellectual review.

References

1. Bartley J. Vitamin D: emerging roles in infection and immunity. Expert Rev Anti Infect Ther. 2010;8(12):1359–1369. doi:10.1586/eri.10.102.
2. Serdar M, Arslan F, Saral N, Yücel D. Correlation between serum 1,25-dihydroxyvitamin D and 25-hydroxyvitamin D in response to analytical procedures: a systematic review and meta-analysis. Turk J Biochem. 2024;49(3):306–324. doi:10.1515/tjb-2023-0258.
3. Office of Dietary Supplements. Vitamin D – Fact Sheet for Health Professionals. National Institutes of Health. (Available from: https://ods.od.nih.gov/factsheets/VitaminD-HealthProfessional/) (Accessed on 07/01/2026)
4. Lamberg-Allardt C, Brustad M, Meyer HE, Steingrímsdóttir L. Vitamin D – Nordic Nutrition Recommendations. Nordic Council of Ministers. (Available from: https://pub.norden.org/nord2023-003/vitamin-d-) (Accessed on 07/01/2026)
5. Demay MB, Pittas AG, Bikle DD, Diab DL, Kiely ME, Lazaretti-Castro M, et al. Vitamin D for the prevention of disease: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2024;109(8):1907–1947. doi:10.1210/clinem/dgae290.
6. Cashman KD, Dowling KG, Škrabáková Z, Gonzalez-Gross M, Valtueña J, De Henauw S, et al. Vitamin D deficiency in Europe: pandemic? Am J Clin Nutr. 2016;103(4):1033–1044. doi:10.3945/ajcn.115.120873.
7. Statistics Canada. Vitamin D status of Canadians, combined cycles, by age group and sex. Table 13-10-0872-01. doi:10.25318/1310087201-eng.
8. Amrein K, Scherkl M, Hoffmann M, Neuwersch-Sommeregger S, Köstenberger M, Tmava-Berisha A, et al. Vitamin D deficiency 2.0: an update on the current status worldwide. Eur J Clin Nutr. 2020;74:1498–1513. doi:10.1038/s41430-020-0558-y.
9. Siddiqee MH, Bhattacharjee B, Siddiqi UR, Rahman MM. High prevalence of vitamin D deficiency among South Asian adults: a systematic review and meta-analysis. BMC Public Health. 2021;21:1823. doi:10.1186/s12889-021-11809-3.
10. Bio-Rad Laboratories. Vitamin D assay sample document. (Available from: https://cdn.bfldr.com/YLD4EVFU/at/scjm7t2858p8thq2gkbwb8/ep28a3ce_sample.pdf) (Accessed on 07/01/2026)
11. Ammer T, Schützenmeister A, Prokosch HU, Rauh M, Rank CM, Zierk J. refineR: a novel algorithm for reference interval estimation from real-world data. Sci Rep. 2021;11:16023. doi:10.1038/s41598-021-95301-2.
12. Jadoon SA, Ahmed A, Alam MA. Vitamin D deficiency in Pakistan: tip of iceberg. J Ayub Med Coll Abbottabad. 2018;30(1):78–80. (Availablr from: https://jamc.ayubmed.edu.pk/index.php/jamc/article/view/3544/1862) (Accessed on 07/01/2026)
13. Roomi MA, Farooq A, Ullah E, Lone KP. Hypovitaminosis D and its association with lifestyle factors. Pak J Med Sci. 2015;31(5):1236–1240. doi:10.12669/pjms.315.7196.
14. Horton-French K, Dunlop E, Lucas RM, Pereira G, Black LJ. Prevalence and predictors of vitamin D deficiency in Australian adolescents and young adults. Eur J Clin Nutr. 2021;75(11):1627–1636. doi:10.1038/s41430-021-00880-y.
15. Forth Life. UK Vitamin D statistics. (Available from: https://www.forthwithlife.co.uk/blog/uk-vitamin-d-statistics/) (Accessed on 07/01/2026)
16. Arshad S, Zaidi SJA. Vitamin D levels among children, adolescents, adults, and elders in Pakistani population: a cross-sectional study. BMC Public Health. 2022;22:2040. doi:10.1186/s12889-022-14526-6.
17. Cui A, Zhang T, Xiao P, Fan Z, Wang H, Zhuang Y. Global and regional prevalence of vitamin D deficiency in population-based studies from 2000 to 2022: a pooled analysis of 7.9 million participants. Front Nutr. 2023;10:1070808. doi:10.3389/fnut.2023.1070808.
18. Ammer T, Schützenmeister A, Prokosch HU, Rauh M, Rank CM, Zierk J. refineR: a novel algorithm for reference interval estimation from real-world data. Sci Rep. 2021;11:16023. doi:10.1038/s41598-021-95301-2.
19. Wierzbicka A, Oczkowicz M. Sex differences in vitamin D metabolism, serum levels and action. Br J Nutr. 2022;128(11):2115–2129. doi:10.1017/S0007114522000149.

20. Usama Z, Rehman Z, Ali A. Establishing reference ranges for vitamin D levels in a representative population of Peshawar, Pakistan. medRxiv. 2024. doi:10.1101/2024.02.15.24302821.
21. Ferrari D, Lombardi G, Banfi G. Concerning the vitamin D reference range: pre-analytical and analytical variability of vitamin D measurement. Biochem Med (Zagreb). 2017;27:030501. doi:10.11613/BM.2017.03050.
22. Perales-Afán JJ, Aparicio-Pelaz D, López-Triguero S, Llorente E, Puente-Lanzarote JJ, Fabre M. Direct and indirect reference intervals of 25-hydroxyvitamin D: it is not a real vitamin D deficiency pandemic. Biochem Med (Zagreb). 2024;34(2):020706. doi:10.11613/BM.2024.020706.

Research Article

Frequency of Rejected Samples in the Hematological Departments: A Comparison Between Palestinian Medical Complex Hospital and Hebron Governmental Hospital

Lubna Mahmoud Habis[1], Mahmoud A. Srour[1,2], Ola Karmi[1*]

[1]*Master's program in Clinical Laboratory Science, Faculty of Pharmacy, Nursing and Health Professions, Birzeit University, Birzeit, Palestine*

[2]*Biology and Biochemistry Department, Faculty of Science, Birzeit University, Birzeit, Palestine*

Article Info

[*]***Corresponding Author:***
Ola Karmi
Master's program in Clinical Laboratory Science, Faculty of Pharmacy, Nursing and Health Birzeit University, Birzeit, Palestine
E-mail: okarmei@staff.birzeit.edu
ORCID ID: 0000-0002-8684-7039

Keywords

Rejected Samples, Preanalytical Errors, Turnaround Time (TAT), Hematology laboratories, Na-citrate blood samples, EDTA blood samples, Laboratory quality

Abstract

Background: Clinical laboratories play a vital role in diagnosis and management of patients. They use established protocols for sample collection and analysis, and several criteria to judge the quality of clinical samples intended for laboratory analysis. Two essential performance measures recommended for clinical laboratories are rejected samples and turnaround time (TAT), which have substantial clinical consequences for patient safety. This study evaluates these two measures in the hematological departments among the two biggest Palestinian Governmental Hospitals in the West Bank region, the Palestinian Medical Complex Hospital (PMCH), and the Hebron Governmental Hospital (HGH). This study is the first one to evaluate the rejected samples from Palestinian governmental hospitals.

Methods: A retrospective analysis of rejected hematology samples was conducted over eight months (February–September 2023). Rejection frequencies, causes, and their influence on TAT were assessed according to the Ministry of Health guidelines.

Results: PMCH exhibited a significantly higher rejection rate (2.1%) than HGH (1%), primarily due to clotted EDTA and Na-citrate tubes. The TAT for rejected samples was longer at HGH (over 10 hours) than at PMCH (5–6 hours).

Conclusion: The study revealed a high specimen rejection rate, and prolonged TAT for EDTA samples, with approximately 40% of rejected samples not repeated within 24 hours. These findings provide baseline data to guide future quality improvement initiatives, staff training, and standardization of preanalytical procedures, ultimately enhancing laboratory quality and patient safety. Larger-scale studies are recommended to further investigate the factors contributing to sample rejection and inform corrective measures.

Introduction

The healthcare system is a growing sector [1]. Medical Laboratory generate the most significant portion of structured data in all medical disciplines [2]. Therefore, it is crucial that all laboratory data reported be accurate, precise, and valid to yield reliable results [3]. When selecting performance measures, it is important to consider those with the greatest impact on patient care, areas where improvements are more feasible, and those that are prone to frequent errors [4].

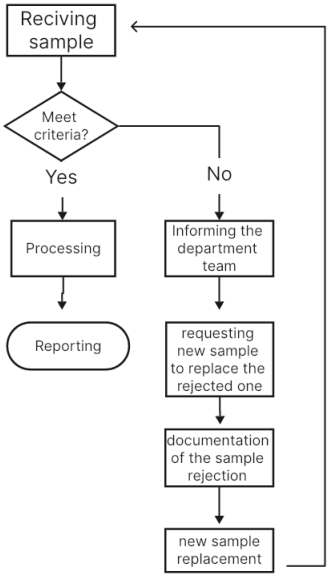
To avoid providing inaccurate results or failed testing, the laboratory should have clear criteria for sample acceptance [5]. Accrediting agencies, such as ISO 15189, require laboratories to implement standardized criteria for sample processing and to enforce a sample rejection system [5, 6]. ISO 15189 accreditation involves standardized procedures for sample collection, handling, and documentation; regular internal audits; staff competency assessments, and validation of analytical methods to ensure reliable results and continuous improvement of laboratory performance [5, 6].

Clinical laboratories worldwide have established protocols for sample collection and analysis, as well as several criteria, to judge the quality of the samples for analysis [7]. Two essential performance measures recommended for clinical laboratories are rejected samples and TAT [8]. Sample rejection has many reasons, which vary between laboratories [9]. Some rejection criteria are common to all laboratories, including: insufficient specimen collection, labeling errors, missing information on the request form, incorrect requests, clotting, inadequate sample volume, improper tubes, hemolysis, and incorrect transportation or storage temperatures [9]. Clotted and hemolyzed samples are the most common causes of sample rejection [9]. These errors are due to inadequate or incorrect knowledge and skills in performing venipuncture, and inadequate training for phlebotomists [10]. Supplementary Table 1 presents the different types of samples and their rejection criteria. Hence, specimen rejection has significant clinical implications for patient care, including the discomfort of re-drawing blood, potential complications such as iatrogenic anemia or hematoma, delay in TAT, and result reporting [11]. Therefore, addressing opportunities in the preanalytical phase is helpful in influencing cost-effectiveness, improving quality, and increasing customer satisfaction [9].

The laboratory technician who receives the sample is responsible for evaluating it according to the established criteria [12]. If a sample is rejected, the laboratory will notify the department where the patient resides. The notification should include the patient's information and reason for rejection. Moreover, this should be recorded clearly, as established by laboratory and hospital policies, which are critical for sample tracing and feedback evaluation [12]. However, the decision to replace the sample or cancel the order should be made by the physician responsible for the patient [12]. (Figure 1) describes the laboratory staff who receive the sample and the process of accepting or rejecting the samples.

TAT indicates the efficiency of the laboratory workflow process [13]. Furthermore, it marks the timeframe for clinicians to plan the management of their patients upon receiving the laboratory results. Hospital Management uses TAT compliance as one of the two yardsticks to evaluate the efficiency of laboratory services [14]. Therefore, this study aimed to evaluate the percentage of rejected samples and TAT in the hematological departments of the two largest Palestinian Governmental Hospitals in the West Bank region. This is the first study in the Palestinian Ministry of Health (PMOH) to systematically assess pre-analytical quality indicators in hematological and routine coagulation testing across two major tertiary governmental hospitals.

Figure 1: Steps of the sample recipient by the laboratory staff and the process of whether accepting or rejecting the samples in hospital laboratories.



Methods

Study Design

This retrospective study was conducted in the Hematology Departments in the laboratories of Palestinian Governmental hospitals. The rejected sample documentation was evaluated, as described in the PMOH regulations. This study was conducted between February 2023 till September 2023. This period was specifically chosen because Hebron Governmental Hospital (HGH) began documenting rejected samples in February 2023. The hospitals' included in this study were the HGH in Hebron, which has 237 beds, provides medical services for more than 700,000 people, and the Palestinian Medical Complex Hospital (PMCH) in Ramallah, which has 312 beds, and provides services for more than one million people. Both hospitals are in the process of being accredited by ISO15189 [6] and follow the same sample acceptance and rejection policies, as established by internal regulations implemented by PMOH.

Data Collection

All data were collected manually from the Avicenna Hospital Information System (HIS) and manual documents. Based on the PMOH criteria for sample rejection, all samples received in the lab and rejected based on the rejection criteria will be documented on the respective form. Some technicians document the rejection reason on HIS, and sometimes will document the rejection on both. This study followed samples rejected in the Hematological department of the hospital laboratories for EDTA samples, which are mainly used for testing Complete Blood Count (CBC), and Na-citrated tubes, which are mainly used for coagulation tests, including Prothrombin Time (PT), and Activated Partial Thromboplastin Time (APTT).

The other parameter evaluated was the TAT. Because the rejected samples will not be reported immediately, and new samples will be requested to be repeated, the value of the TAT will be evaluated for two time points. The sample TAT of the results from the first request time until the final report is referred to as True TAT. The sample TAT, extending from the receipt of the new sample until the final reporting is referred to as Actual TAT.

Data Analysis and Statistics

Results were presented using the mean percentage from each month of rejected samples per month, then the graphs were plotted using the Mean ± Standard Deviation (SD) of the evaluated period. Significant differences between groups were evaluated using the student's independent t-test and *P-value*, in which significance was considered as $P^{*}<0.05$, $P^{**}<0.01$, and $P^{***}<0.001$.

Ethics Approval and Consent

The study protocol was approved by the Birzeit University Scientific Research Committee. Then approved by the Palestinian Ministry of Health, the Medical Laboratories Department (No. 2024-353). All patients' data were anonymized and handled with high confidentiality.

Results

Evaluating the Rejected Samples Percentage and Documentation in the Hematology Departments in HGH and PMCH

The samples were collected from various departments in both hospitals, between February and September in the year 2023. The total number of blood samples received by the hematology departments in the HGH hospital laboratory during the study period (8 months) were 61,254 samples, and at the PMCH hospital laboratory, there were 56,243 samples, as presented in (Supplementary Figure 1A).

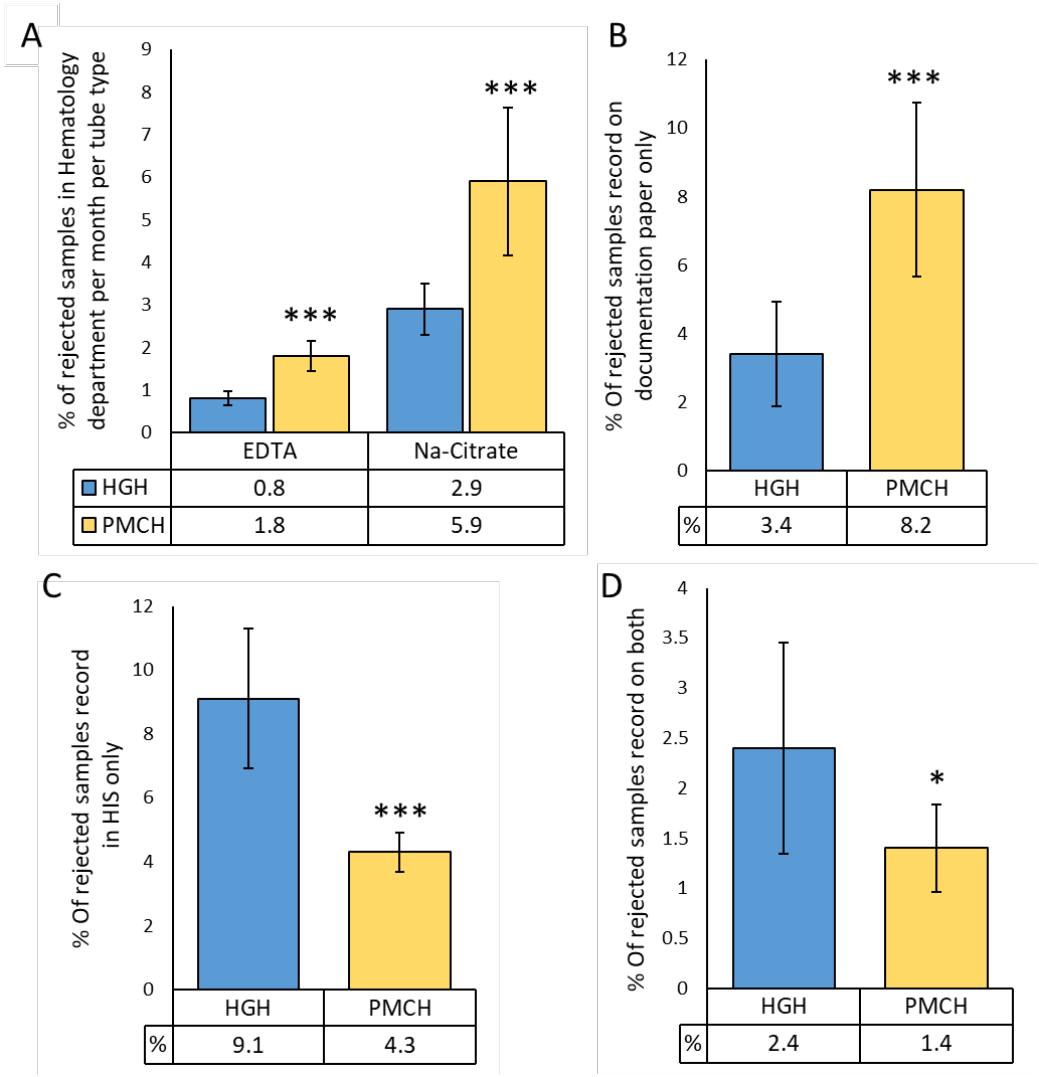
From these samples received in the hematology departments, 640 samples were rejected at the pre-analytical stage and weren't processed in HGH, and 1,157 in PMCH (Supplementary Figure 1B). The percentage of rejected samples in these hospitals was 1% ± 0.18 per month at the HGH, while it was 2.1% ± 0.42 per month at the PMCH (Supplementary Figure 1C). By distributing the samples depending on the tube type of EDTA tubes or Na-Citrated tubes, it was found that the mean percentage of rejected EDTA samples was 1.8% per month at PMCH compared to 0.8% per month at HGH (Figure 2A).

The mean percentage of rejected Na-citrated samples was 5.9% per month at the PMCH, compared to 2.9% per month at the HGH (Figure 2A).

By investigating the distribution of the percentage of the rejected samples by referring to the wards at the hospitals, the highest mean percentage of the EDTA and the Na-citrated tubes was at the PMCH hospital, as presented in (Supplementary Figure 2). The main criteria for sample rejection in both the PMCH and the HGH were Clotted, Insufficient (had less volume than needed), Hemolyzed, Incorrect sample, Overfilled samples, Diluted blood samples with intravenous fluids, and samples collected from patients and received in the laboratory without a request for laboratory testing. At the PMCH and HGH laboratories, the main rejection criteria were the clotted samples in the EDTA tubes, with a percentage of 11.2% and 11.9% per month, respectively, as presented in (Supplementary Figure 3).

Tracking the rejected samples' documentation throughout the study period (Figure 2B, C, D), revealed an inconsistency in the documentation process. Part of the samples were documented properly on the rejected files manually, with a percentage of 8.2% at the PMCH and 3.4% at the HGH (Figure 2B), while rejected samples documented only on the HIS software were of 4.3% at the PMCH and 9.1% at the HGH (Figure 2B). On the other hand, rejected samples were documented by both systems (paper and HIS) in 1.4% and 2.4% of the rejected samples at PMCH and HGH, respectively (Figure 2C).

Figure 2: Evaluating the Rejected Samples Percentage and Documentation in the Hematology departments in HGH and PMCH.

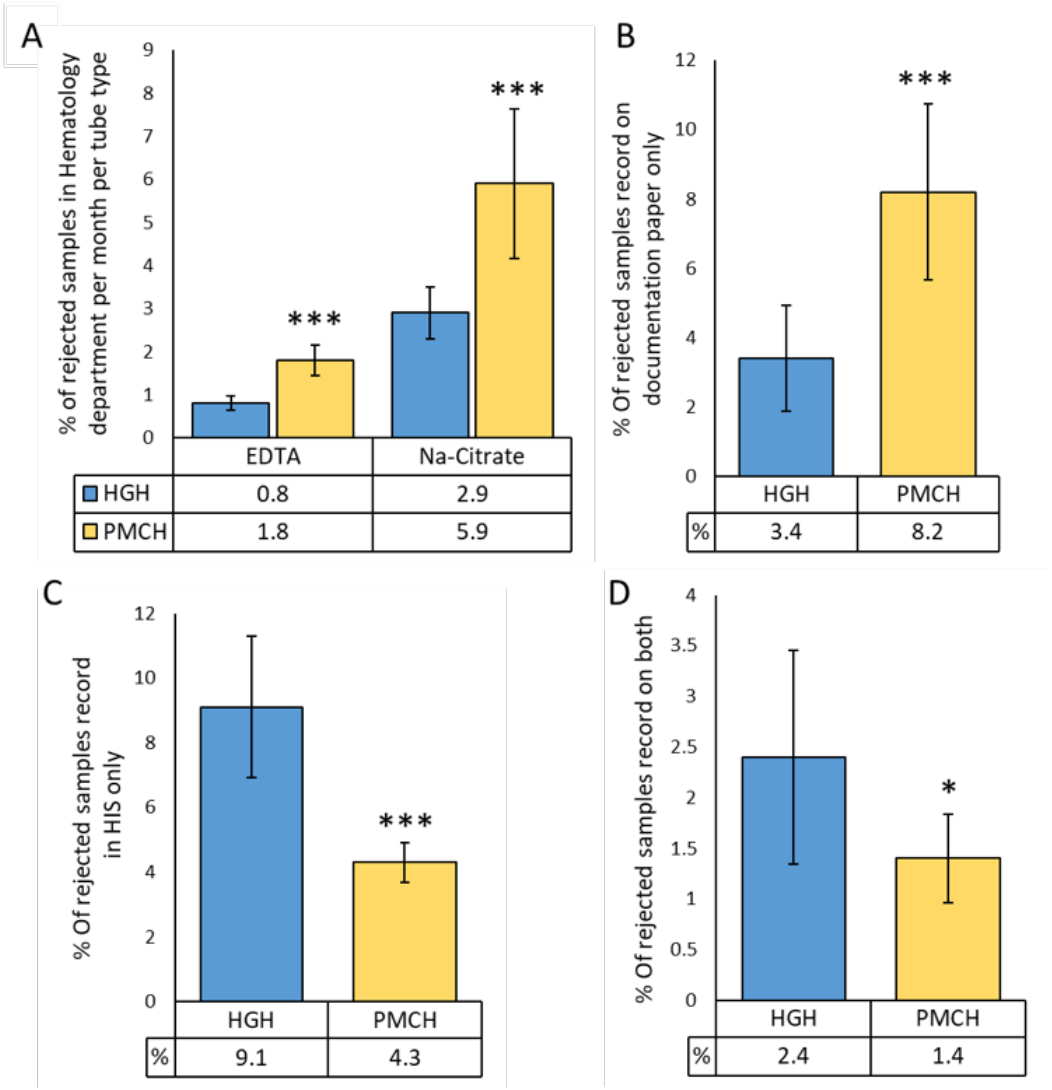


*(A) Comparison of the mean percentage of the total rejected EDTA tubes in the Hematology departments of both hospitals and the comparison of the mean percentage of the rejected Na-Citrated samples in the Hematology departments of both hospitals per month in the period Feb –Sep 2023. (B) Frequency of the rejected samples documentation using the manual paper documentation method in both hospitals, the HGH (blue column) and the PMCH (yellow column). (C) Frequency of the rejected samples documentation on the HIS software in both hospitals. (D) Documentation of the rejected samples on both the paper documents and the HIS software at the same time, in both hospitals. Results are represented as mean±SD per month of the study period. Significant difference was evaluated using independent T. Test, * P < 0.05, *** P < 0.001.*

Evaluating the TAT for the Rejected Samples in PMCH and HGH, Concerning the Critical Values

The average of the TAT for the EDTA tubes or the Na-citrated samples at the PMCH was significantly lower than that at the HGH, where the true TAT were 6 hours and 52 minutes at the PMCH and 10 hours and 38 minutes at the HGH. While for the Na-citrated samples, the average TAT was 5 hours and 1 minute for the PMCH and 10 hours and 59 minutes at the HGH (Figure 3A). On the other hand, the Actual TAT of the EDTA tube samples was 27 minutes at the PMCH and 19 minutes at the HGH, but for the Na-citrated samples, the average TAT was 34 minutes at the PMCH and 41 minutes at the HGH (Figure 3B), the results showed no statistically significant difference. Evaluation of the critical values' TAT is an essential step in evaluating the efficiency of the laboratory. Investigating the values reported for the rejected samples after resending them showed that the percentage of rejected samples reported as critical values was 7.8% in the PMCH for the EDTA tubes and 3.1% for the HGH (Figure 3C). The critical values for the Na-citrated samples were less than those for the EDTA, since for the PMCH it was 0%, but for the HGH it was 0.6%, (Figure 3C). By investigating the time needed to report these critical values, we calculated the true TAT for both hospitals and showed that both were close in these results with almost 6 hours for both of them, while the actual TAT was less at the HGH (15 minutes), compared to the PMCH (1 hour and 37 minutes) (Figure 3D), the results showed no statistically significant difference.

Figure 3: Comparison of the TAT for the rejected samples in PMCH and HGH, concerning the critical values.

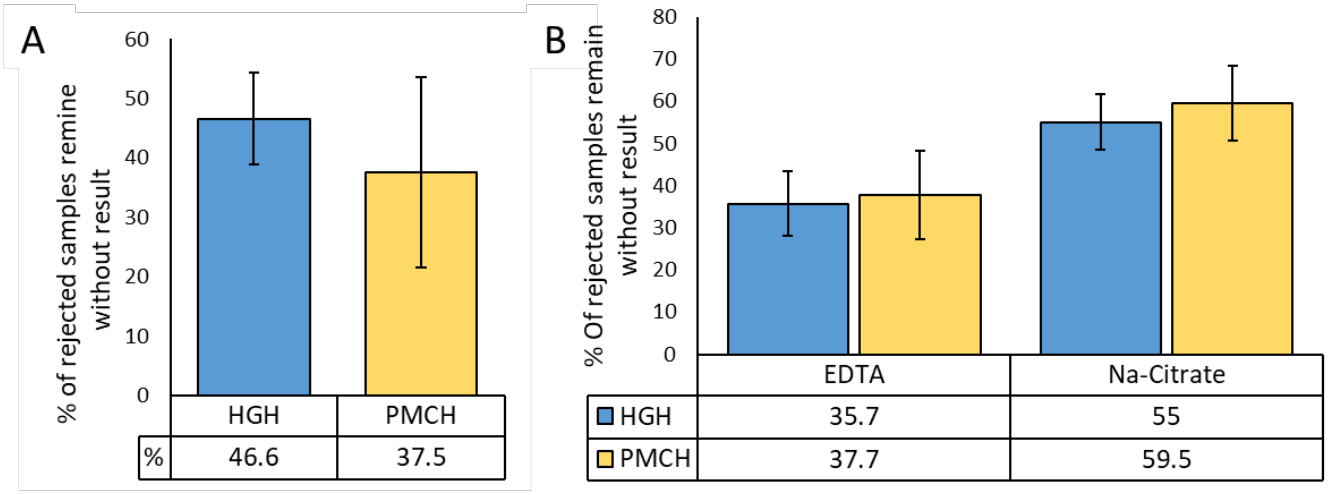


*(A) True TAT average of the rejected samples of the EDTA samples and the Na-Citrated samples in the PMCH and the HGH. The blue columns represent HGH, and the yellow columns represent PMCH. (B) Actual TAT average of the rejected samples of the EDTA samples and the Na-Citrated samples in the PMCH and the HGH. (C) Critical value frequency for the rejected samples upon being repeated, for EDTA and Na-citrated samples in PMCH and HGH. (D) True TAT for critical values reported after repeating the rejected samples by calculating the True and Actual TAT, at the PMCH and HGH. *P < 0.05, **P < 0.01.*

Rejected Samples but not Repeated Within 24 Hours

The most direct consequence of specimen rejection is the need to collect a new specimen for the patient to replace the rejected sample. Recollecting samples from patients is challenging; sometimes patients refuse to give blood samples again, or they have been discharged from the hospital, or even the healthcare person in charge will not collect another sample for any reason. As a result, these patients receive no results for the originally ordered tests. By comparing the hospitals enrolled in the study, the PMCH and the HGH, both had almost similar frequencies for no results for repeated samples. In total, 37.6% of rejected samples remained without results in PMCH, meaning that no second samples were prepared to replace the rejected ones, while this frequency was 46.6% in the HGH (Figure 4A). Further analysis of rejected samples that remain without results revealed that no significant difference was observed for EDTA and Na-citrated samples in both PCMH and HGH (Figure 4B).

Figure 4: Rejected Samples without repeating sampling within 24 hours.



(A) The graph represents the frequency of rejected samples, that have no reported results within 24 hours, the blue columns represent the HGH, and the yellow columns represent the PMCH. (B) Graph represented the frequency of the rejected samples of the EDTA or Na-citrate samples, comparison between the PMCH and HGH.

Discussion

Quality control (QC) monitors the total testing process to consider and identify laboratory errors, which ensures accurate and reliable results and increases confidence in laboratory performance [15]. Institutions that adopt patient safety measures, and implement corrective interventions, can make healthcare safer for patients and healthcare workers. This can happen by identifying the factors and events contribute to medical errors. Developing multifaceted prevention protocols and implementing these strategies at various healthcare levels [16].

Recently, research highlighted that one of the most important QC measures in medical laboratories is the rejection of samples. These lead to delays in sample analysis and prevent clinicians from providing timely management to their patients, which negatively impacts patients' outcomes [17]. In our study, we chose the PMCH in Ramallah, and the HGH in Hebron. Rejected samples' records in the hematology laboratory were evaluated from different perspectives to understand the main sources of errors, as well as to understand the possible correction measures that should be taken. This study revealed a significant number of rejected samples, which are likely to result in delays in diagnosing and treating patients.

The results were compared to other countries; a recent comprehensive meta-analysis study reported an average percentage of blood sample rejection of 1.99% [9]. Another study from Malaysia reported a high sample rejection rate of 3.38% in the hematology department [17]. While recent reports from Saudi Arabia showed a sample rejection rate of 1.3% in the hematology department [18]. Correspondingly, the rejected sample rate in the Palestinian MOH-related hospitals is considered similar to the average percentages reported in other countries. This reflects the increased quality of follow-ups of the system. Alternatively, this decrease in rate can be explained by the lack of documentation of the samples, which may need more evaluation and follow-up within the same hospitals and an audit from the MOH. By investigating specific tests based on EDTA tube, the rejection rate for this sample was 0.8% at HGH and 1.8% at PMCH, which are significantly lower compared to studies from Lahore, and Sri Lanka, where the rejection rates were 5.15% and 3.3%, respectively [19, 20]. Additionally, Na–citrated samples exhibited the highest percentage of rejection rate in the Hematology department of the investigated Palestinian hospitals. Indeed, coagulation samples are sensitive due to the need for a precise anticoagulant-to-blood ratio, and timely processing to prevent clotting or degradation [21]. Several factors contribute to the variation in rejection rates, such as type and duration of the study, sample size, rejection criteria between hospitals, laboratory personnel awareness, and samples received from different wards [12, 22].

By looking at the most common reason to reject the samples, clotted samples is the highest reason in both hospitals, which reflects a possible lack of knowledge of the phlebotomists, according to the instructions provided by the tubes' manufacturer [23]. Similar findings have been reported in various studies worldwide, including those conducted in China, Korea, India, and Malaysia, which report that the improper mixing of the sample was the reason for blood clotting after sample collection [12, 20, 24, 25]. Insufficient samples or under-filling of tubes is the second reason for sample rejection in both HGH and PMCH in the Hematology department, consistent with the findings from the mentioned countries [12, 20, 24, 25]. In contrast, a study done at Chughtai Lab, Lahore [19], found that labeling errors and storage errors were

considered the most significant reasons for rejection, followed by diluted samples [19]. The higher frequency of insufficient samples refers to the difficulty in collecting sufficient blood samples from newborns, children, and ICU patients [9, 17]. Although hemolyzed samples, though not recorded in rejected EDTA samples at HGH, remain a significant portion of all rejected samples. To address this, hospitals should implement internal training for phlebotomy staff [12]. Errors in test ordering, where samples are received without a physician's request, also contribute to rejections. Improved communication and cooperation between medical teams need to be addressed [12]. At HGH, the Na-Citrate tube had more "no order" cases compared to EDTA samples, likely due to unfamiliar test ordering. The occurrence of 'no order' cases is linked to inadequate inspection of sample collection forms and a shortage of personnel relative to patient demand for laboratory services [12].

Considering the complexity of the healthcare system, documentation of records is an essential component of communication [26]. The documentation can be paper-based, electronic, or both. Organizations are required to establish a clear system to track and follow patient information [27]. On the other hand, organizations should have the ability to follow up and meet the respective standards [28]. Rejected samples criteria at the PMOH hospitals are requested to be documented using the paper system and per the respective Standard Operating Procedures (SOP). By following up on the MOH-related hospitals' records, it was obvious that there is no consistency in the documentation, which indicated miscommunication between the laboratory employees and the documentation system established by the hospital based MOH guidelines. There was rejection data documented using the paper documents, or the HIS, or both. This inconsistent documentation decreases the confidence and reliability of the whole documentation and recording of the rejection sample, as well as for any other documents that include mistakes [3]. These mistakes and loss of credibility jeopardize patient safety and the ability to follow up, which affects medical decision-making [3]. However, reporting this inconsistency, it gives the possibility that part of the rejected samples was not reported at all. This should be evaluated in future evaluations of the rejection system by following up on the rejection system daily and spontaneously. From our perspective, the documentation systems and hospital policies can be updated and improved to reflect the hospital's needs [13].

Sample rejection prevents sample analysis, which increases TAT and causes delays in patient diagnosis and treatment [9]. The rejection of samples impedes the analysis process and necessitates new sample requests [14]. Moreover, timely reporting of trustworthy results, especially critical values save the patient's life, and supports effective clinical decisions [29]. By evaluating the TAT in the PMCH and HGH, the results of the critical values were taking a very long time to be reported since the first request time, which we referred to as the True TAT, but it was obvious that the action to replace the samples at the PMCH was more efficient than in the HGH, that was taking almost twice the time needed to be reported in PMCH. The percentage of critical values in the EDTA rejected samples criteria was significantly higher at the PMCH, which may need more attention and orientation for the medical staff responsible for sample collection. In our study, the explanation for the long true TAT may be due to the greater complexity of very large hospitals, which could contribute to less effective communication among healthcare staff members [30].

A significant consequence of laboratory sample rejection is the state of "no results" available for the rejected samples. Explanations for this include discharge of the patients before repeated sample collection, or inappropriate communication between the health work team. At the governmental hospitals, laboratory testing costs are reimbursed through governmental insurance in some cases or by the patients themselves; these costs are supported by the MOH, so the payments are not considered high. Even though some patients may refuse repetitive payments. Also, patients' rights are to refuse the recollection of blood samples, which may jeopardize the patient's safety and privacy; it is much better to get the sample properly the first time [31]. On the other hand, "no results" of these requests highlight the probability of system abuse. Questioning the reality of the actual need for blood collection from the beginning. There is a probability of wasting time for blood collection and wasting time for laboratory analysis, additionally, exposing the system to consuming unnecessary and expensive resources [15].

This study has several limitations; it was conducted only in two laboratories in governmental hospitals, and only in the hematology department. The research was done for the first time in the Palestinian health system. The quality recording system in the laboratories of the PMOH was implemented a few months before this study was performed, and the new regulations were determined to be present by laboratory technicians. However, it is retrospective and depends on the accuracy of documented records; differences in technician experience and adherence to sample collection procedures may have influenced rejection rates. The study aimed to document the present situation for corrective action measures to be implemented in the future.

Conclusion

Although this study was conducted in two Palestinian governmental hospitals, its findings have impact on laboratories worldwide. Pre-analytical errors are mostly preventable. As our findings showed, the main reason for sample rejection was insufficient communication between the healthcare team and inadequate training. The high TAT for the rejected samples is considered dangerous, especially when critical samples are reported for these samples. So, documentation of the rejected samples monthly, in addition to organizing periodic training for new healthcare personnel, will enhance the reliability of

laboratory test samples and results, which prevents TAT delays and consequently improves patients' outcomes. Advanced training programs for laboratory employees and information about the methods of recording should be clear enough for laboratory technicians. Policies and regulations can be modified as needed through the laboratory employees' experience and practice. Finally, modern technology usage, for keeping documents, will be useful for tracing and highlighting errors and mistakes, which will provide fast and efficient intervention for correction, as well as a high level of analysis to provide evaluation of the system quality.

Acknowledgment
The authors wish to acknowledge the late Dr. Osama Najjar; additionally, we acknowledge the Palestinian Ministry of Health (PMOH) and all contributors to the work in PMOH.

Declaration of competing interest
The authors declare no conflict of interest.

Data Availability
All study data are included in the article and/or supporting information.

Funding Acknowledgment
Master's program in Clinical Laboratory Science, Faculty of Pharmacy, Nursing and Health Professions, Birzeit University, Birzeit, Palestine.

Ethics approval statement
The study protocol was approved by the Birzeit University Scientific Research Committee. The Palestinian Ministry of Health, Medical Laboratories Department (No. 2024-353).

References

1. Durrani H. Healthcare and healthcare systems: inspiring progress and future prospects. Mhealth. 2016;2:3. doi:10.21037/mhealth.2016.01.01.
2. Hulsen T, Friedecký D, Renz H, Melis E, Vermeersch P, Fernandez-Calle P. From big data to better patient outcomes. Clinical Chemistry and Laboratory Medicine (CCLM). 2023;61(4):580–586. doi:10.1515/cclm-2022-0872.
3. Gunnur Dikmen Z, Pinar A, Akbiyik F. Specimen rejection in laboratory medicine: Necessary for patient safety? Biochemia Medica. 2015;25(3):377–385. doi:10.11613/BM.2015.037.
4. Plebani M, Aita A. Patient safety in laboratory medicine. Textbook of patient safety and clinical risk management. 2021. doi:10.1007/978-3-030-59403-9_30.
5. Magwai T, Warasally Z, Naidoo N, Gounden V. Reducing sample rejection in Durban, South Africa. Clinical Chemistry and Laboratory Medicine (CCLM). 2021;59(4):687–692. doi:10.1515/cclm-2020-0974.
6. Plebani M, Sciacovelli L. ISO 15189 accreditation: navigation between quality management and patient safety. Journal of Medical Biochemistry. 2017;36(3):225–231. doi:10.1515/jomb-2017-0019.
7. Mosha VV, Kabanyana C. The rate of sample rejection and pre-analytical errors at KCMC Clinical Laboratory in Moshi, Kilimanjaro. East Africa Science. 2021;3(1):118–124. doi: Not available.
8. Howanitz PJ. Errors in laboratory medicine: practical lessons to improve patient safety. Archives of Pathology & Laboratory Medicine. 2005;129(10):1252–1261. doi:10.1043/1543-2165(2005)129[1252:EILMPL]2.0.CO;2.
9. Getawa S, Aynalem M, Melku M, Adane T. Blood specimen rejection rate in clinical laboratory: A systematic review and meta-analysis. Practical Laboratory Medicine. 2023;33: e00303. doi: 10.1016/j.plabm.2023.e00303.
10. Bölenius K, Söderberg J, Hultdin J, Lindkvist M, Brulin C, Grankvist K. Minor improvement of venous blood specimen collection practices in primary health care after a large-scale educational intervention. Clinical Chemistry and Laboratory Medicine (CCLM). 2013;51(2):303–310. doi:10.1515/cclm-2012-0423.
11. Tesfaw HM, Segaye A, Hassen F. Frequency of specimen rejection and associated factors at St. Paul's Hospital Millennium Medical College, Addis Ababa Ethiopia. Journal of Multidisciplinary Research in Healthcare. 2015;2(1):1–16. doi:10.15415/jmrh.2015.21001.
12. Khang A, Rani S, Sivaraman AK. AI-centric smart city ecosystems: technologies, design, and implementation. CRC Press; 2022. doi:10.1201/9781003205967.
13. Bhati D, Deogade MS, Kanyal D. Improving patient outcomes through effective hospital administration: a comprehensive review. Cureus. 2023;15(10). doi:10.7759/cureus.38139.
14. Yadav C, Saxena R, Kamini V, Harith AK. Can turnaround time compliance in hospital laboratories be set to auto pilot mode? –A 2 years' experience. 2022. doi: https://doi.org/10.9734/jammr/2022/v34i2231582.
15. Mrazek C, Lippi G, Keppel MH, Felder TK, Oberkofler H, Haschke-Becher E, et al. Errors within the total laboratory testing process. Biochemia Medica. 2020;30(2):215–233. doi:10.11613/BM.2020.020502.
16. Rodziewicz TL, Houseman B, Hipskind JE. Medical error reduction and prevention. 2018. doi: https://www.ncbi.nlm.nih.gov/books/NBK499956/.
17. Kamal F, WA WMS, Adbul Monir M. Analysis of sample rejection and the impact on quality of care in patients in a single tertiary healthcare facility in Malaysia. Environmental Proceedings Journal. 2020;5(13):175–181. doi:10.21834/eproceedings.v5i13.2186.
18. Iqbal MS, Tabassum A, Arbaeen AF, Qasem AH, Elshemi AG, Almasmoum H. Preanalytical errors in a hematology laboratory: an experience from a tertiary

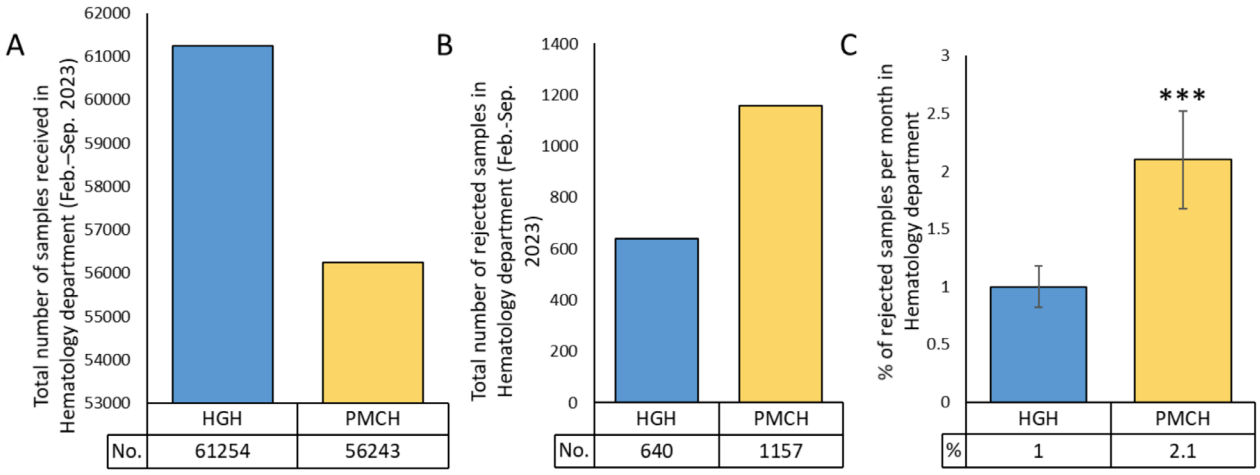
care center. Diagnostics. 2023;13(4):591. doi:10.3390/diagnostics13040591.
19. Noor T, Imran A, Raza H, Umer S, Malik NA, Chughtai AS, et al. An overview of complete blood count sample rejection rates. Cureus. 2023;15(1). doi:10.7759/cureus.33267.
20. Dilshika L, Bandara W, Karunanayaka A. A study on sample rejection rates due to pre-analytical errors. 2020. doi:10.31357/jhsir.v1i01.4771.
21. Bonar R, Favaloro EJ, Adcock DM. Quality in coagulation and haemostasis testing. Biochemia Medica. 2010;20(2):184–199. doi:10.11613/BM.2010.023
22. Mesganaw B, Hassen F, Molla H, Misganaw K. Laboratory specimen rejection rate and associated factors among referred specimens at Debre Markos Referral Hospital, Ethiopia: prospective cross-sectional study. Pan African Medical Journal. 2024;47(1). doi:10.11604/pamj.2024.47.1.39331.
23. Ye Y, Wang W, Zhao H, He F, Zhong K, Yuan S, et al. Haematology specimen acceptability: a national survey in Chinese laboratories. Biochemia Medica. 2018;28(3):420–429. doi:10.11613/BM.2018.030703.
24. Narang V, Kaur H, Selhi PK, Sood N, Singh A. Preanalytical errors in hematology laboratory- an avoidable incompetence. Iranian Journal of Pathology. 2016;11(2):151–154. doi:10.30699/ijp.2016.151.
25. Thate JA, Couture B, Schnock KO, Rossetti SC. Information needs and the use of documentation to support collaborative decision-making: implications for the reduction of central line–associated bloodstream infections. CIN: Computers, Informatics, Nursing. 2021;39(4):208–214. doi:10.1097/CIN.0000000000000703.
26. Stausberg J, Koch D, Ingenerf J, Betzler M. Comparing paper-based with electronic patient records: lessons learned during a study on diagnosis and procedure codes. Journal of the American Medical Informatics Association. 2003;10(5):470–477. doi:10.1197/jamia.M1308.
27. Demsash AW, Kassie SY, Dubale AT, Chereka AA, Ngusie HS, Hunde MK, et al. Health professionals' routine practice documentation and its associated factors in a resource-limited setting: a cross-sectional study. BMJ Health & Care Informatics. 2023;30(1):e100699. doi:10.1136/bmjhci-2022-100699.
28. Li R, Wang T, Gong L, Dong J, Xiao N, Yang X, et al. Enhance the effectiveness of clinical laboratory critical values initiative notification by implementing a closed-loop system: a five-year retrospective observational study. Journal of Clinical Laboratory Analysis. 2020;34(2):e23038. doi:10.1002/jcla.23038.
29. Sun S-CP, Garcia J, Hayden JA. Repeating critical hematology and coagulation values wastes resources, lengthens turnaround time, and delays clinical action. American Journal of Clinical Pathology. 2018;149(3):247–252. doi:10.1093/ajcp/aqx157.
30. World Health Organization. WHO guidelines on drawing blood: best practices in phlebotomy. Geneva: World Health Organization; 2010. doi: ISBN-13: 978-92-4-159922-1.

Supplementary files

Supplementary Table 1: Rejected samples criteria in the clinical laboratory department.

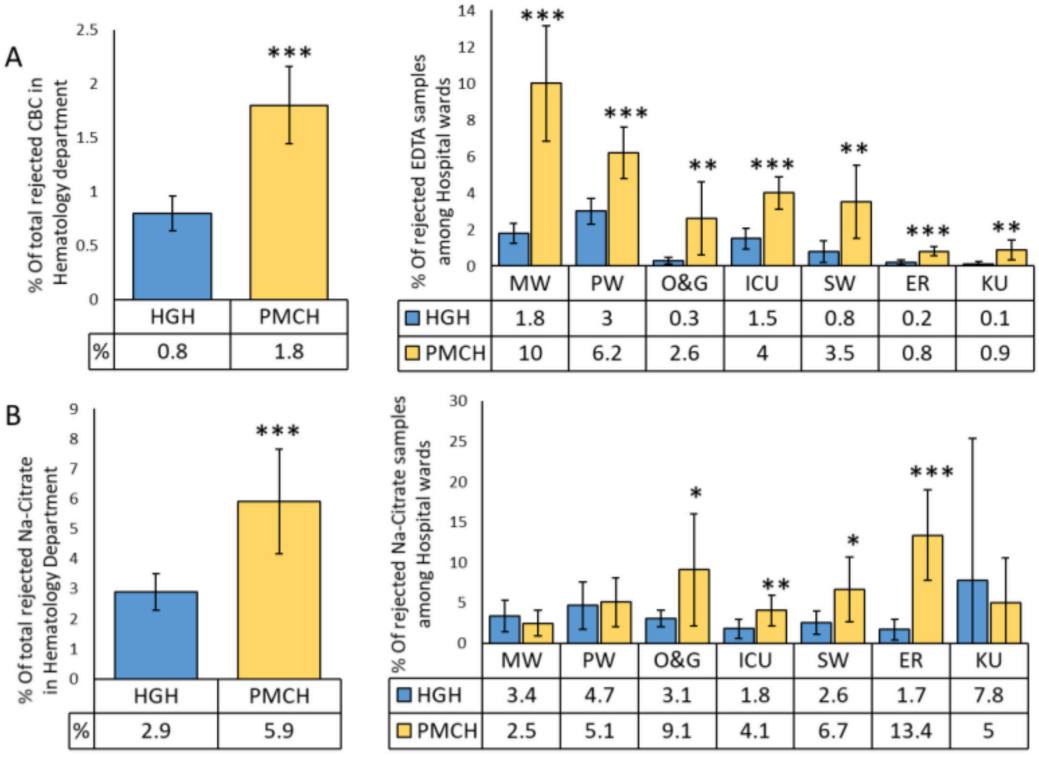
Clinical Laboratory section	Sample Type	Rejection criteria	The most common causes of rejection	Consequences of the problem
Hematology (Lavender top)	EDTA	Clotted sample	1. Inappropriate mixing of the tubes. 2. Lack of phlebotomy training [1].	1. Pseudo-thrombocytopenia. 2. False leukopenia, aberrant red cell indices. 3. Instrument downtime due to probe clogging [2].
		Insufficient samples	1. Incorrect phlebotomy techniques. 2. Difficult venous access [1].	1. Erroneous laboratory results due to improper anticoagulant-to-blood ratio. 2. Redraws [2].
		Excessive amount	1. Clot due to insufficient anticoagulant [2].	1. Erroneous laboratory results [2].
		Labeling errors	1. Missing data. 2. The specimen container labeling away from the specimen collection site [2].	1. Incorrect or delayed diagnosis. 2. Incorrect Treatment [3].
		Lipemic samples	1. Non-fasting patient. 2. Pre-existing metabolic disorders [4].	1. Interference of fat with the optical reading of an instrument leads to incorrect result reporting [4].
		Hemolysis samples	1. Use of inappropriate needle 2. Prolonged tourniquet application 3. Evacuation of the syringe into a tube [1].	1. Interference with spectrophotometric assays [5].
		Contaminated sample	1. Collection by intravenous catheters [6]. 2. Incorrect tube order during phlebotomy. 3. Contamination with alcohol [7].	1. Falsely Pancytopenia [2].
Coagulation (Blue top)	Citrated	Clotted sample	1. Overfilling causes insufficient anticoagulants. 2. Inappropriate mixing of samples [8].	1. Erroneous laboratory results [2].
		Labelling Error	1. Missing data. 2. Label the specimen container away from the specimen collection site [2].	1. Incorrect or delayed diagnosis. 2. Incorrect Treatment [3].
		Overfilled Tube	1. Clot due to insufficient anticoagulant. 2. Direct free flow dripping of blood into the tube in pediatric [8].	1. Erroneous Coagulation test results [9].
		Collected Via an indwelling catheter	1. Improper Anticoagulant-to-blood ratio [10].	1. Erroneous laboratory results [10].
		Hemolysis sample	1. Rough handling of the samples. 2. Vigorous shaking 3. Prolonged centrifugation. 4. Use of inappropriate needle size. 5. Prolong tourniquet application 6. Evacuation of the syringe into a tube [2].	1. Significant analytical interference [5].

Supplementary Figure 1: Comparison of the distribution of the rejected samples in the Hematology.



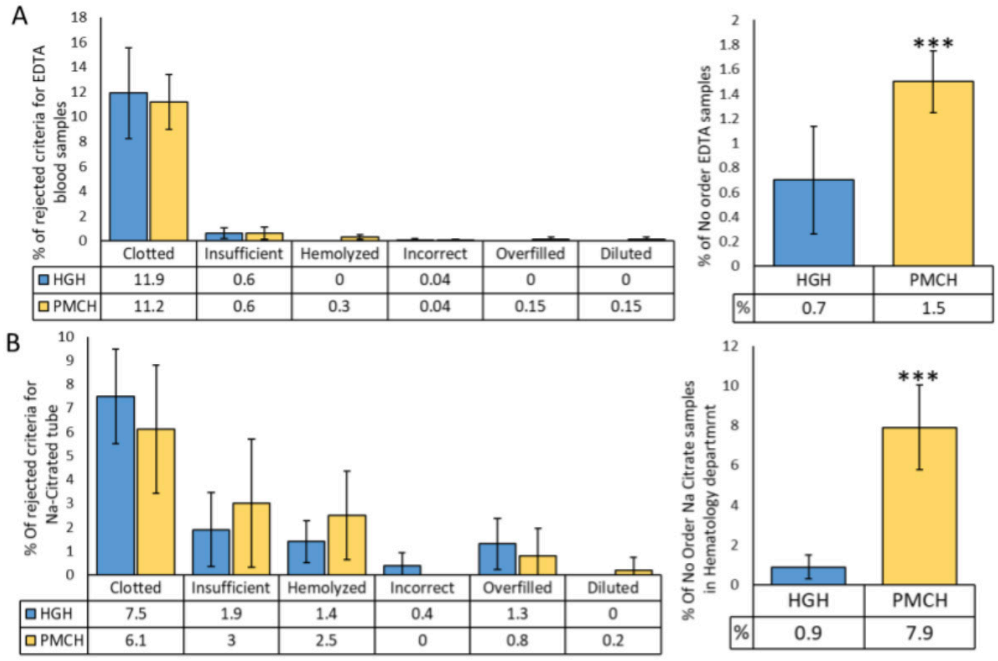
*(A) Total number of samples received at the Laboratory's Hematology department at the Hebron Governmental Hospital (HGH) and the Palestine Medical Complex Hospital (PMCH) during the study period. The blue column represents the HGH, while the yellow column represents the PMCH. (B) Total number of rejected samples in both hospitals during the study period. (C) The mean percentage of the rejected samples in both hospitals per month. Results are represented as mean (of the monthly percentage) ± SD during the months of the study period. Significant difference was evaluated using an independent t-test, *** P < 0.001.*

Supplementary Figure 2: Frequency of the EDTA and the Na-citrated blood rejected samples, referred to the requesting department in HGH and the PMCH hospitals, from February to September 2023.



*(A) The percentage of rejected EDTA samples in the hematology departments, by referring to the percentage distribution of these samples to the blood test requesting word at the hospitals. HGH are presented by blue columns and the PMCH are presented by yellow columns (B) The percentage of the rejected Na-Citrate samples in the hematology departments, by referring to the percentage distribution of these samples to the blood test requesting word at the hospitals, throughout the study. The analyzed departments included the Medical Ward (MW), Pediatric Ward (PW), Obstetrics and Gynecology (O&G), Intensive Care Unit (ICU), Surgical Ward (SW), Emergency Department (ED), and Kidney Unit (KU). Results are represented as mean ± SD per month of the study period. Significant difference was evaluated using independent T. Test, *P < 0.05, **P < 0.01, ***P < 0.001.*

Supplementary Figure 3: Rejection criteria distribution for EDTA and Na-Citrated blood samples in HGH and PMCH.



*(A) (Left Panel) Graph displays the frequency of each rejection criterion for EDTA samples across the PMCH and the HGH departments, within the study period. (Right Panel) Graph presenting the percentage of the EDTA samples received to the HGH and the PMCH laboratories with no order per month. (B) (Left Panel) The graph presents the frequency of each rejection criterion for Na-Citrated samples across the HGH and the PMCH laboratories. The blue columns represent the HGH, while the yellow columns represent the PMCH. The rejection criteria include clotted samples, insufficient samples, hemolysis, incorrect labeling, overfilled samples, and diluted samples. Results are represented as mean ± SD per the months of the study period. Significant difference was evaluated using an independent T. Test, *** $P < 0.001$.*

Supplementary References

1. Pozniak, Y., et al., System-wide clinical proteomics of breast cancer reveals global remodeling of tissue homeostasis. Cell systems, 2016, 2(3):172-184.
2. Tyanova, S., et al., The Perseus computational platform for comprehensive analysis of (prote) omics data. Nature methods, 2016, 13(9):731.
3. Gunnur Dikmen, Z., A. Pinar, and F. Akbiyik, Specimen rejection in laboratory medicine: Necessary for patient safety? Biochemia Medica, 2015, 25(3):377-385.
4. Cox, J., et al., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics, 2014, 13(9): 2513-2526.
5. Sasi, U., et al., Mitochondria associated membranes (MAMs): emerging drug targets for diabetes. Current Medicinal Chemistry, 2020.
6. Costello, J.C., et al., A community effort to assess and improve drug sensitivity prediction algorithms. Nature biotechnology, 2014, 32(12):1202.
7. Vallese, F., et al., ER-Mitochondria Calcium Transfer, Organelle Contacts and Neurodegenerative Diseases, in Calcium Signaling. 2020,Springer: 719-746.
8. Torti, S.V. and F.M. Torti, Iron and cancer: more ore to be mined. Nature Reviews Cancer, 2013, 13(5):342.
9. Tong, L., et al., Reactive oxygen species in redox cancer therapy. Cancer letters, 2015, 367(1): 18-25.
10. Poillet-Perez, L., et al., Interplay between ROS and autophagy in cancer cells, from tumor initiation to cancer therapy. Redox biology, 2015,4:184-192.
11. Sullivan, L.B. and N.S. Chandel, Mitochondrial reactive oxygen species and cancer. Cancer & metabolism, 2014,2(1):17.

Research Article

Comparison of the glucose stability effect between Citrate–sodium fluoride–EDTA, Heparin, and Sodium fluoride containing tubes: a prospective experimental study in the Biochemistry Laboratory at the University Clinic of Kinshasa, DR Congo

Ongona Sheko Nounou[1], Ngole Mamy[*1,2,3], Pambu Dahlia[1], Tshibuela Beya Dophie[1], Beya Michael[4] , Mukadi Patrick[4,5]

[1]*Department of Medical Biology, University Clinics of Kinshasa, DR Congo*
[2]*Center for Human Genetics, University of Kinshasa, DR Congo*
[3]*Reference Center for Rare and Undiagnosed Diseases, University of Kinshasa, DR Congo*
[4]*Institute National de Recherches Biomedicales" (INRB), Kinshasa, DR Congo*
[5]*Microbiology Laboratory, Department of Medical Biology, University Clinics of Kinshasa, DR Congo*

Article Info

[*]***Corresponding Author:***
Mamy Ngole
Department of Medical Biology, University Clinics of Kinshasa; Center for Human Genetics, University of Kinshasa; Reference Center for Rare and Undiagnosed Diseases, University of Kinshasa, DR Congo
E-mail: mamy.ngole@unikin.ac.cd
ORCID: 0000-0002-1041-9257

Keywords

diabetes mellitus, glucose stability, glycolysis inhibitors, pre-analytical glucose

Abstract

Tubes containing sodium fluoride (NaF) are no longer recommended for measuring blood glucose levels. This prospective experimental study was conducted at the Biochemistry Laboratory of the University Clinics of Kinshasa,in the Democratic Republic of the Congo, and included seventy healthy volunteer blood donors. Glucose stability was assessed in tubes containing NaF,lithium heparinised tubes, and a tube containing a mixture of citrate,NaF and EDTA.The tubes containing citrate and NaF were stored at room temperature,whilst the corresponding heparinised tubes were placed on ice. Glucose measurements were taken at various time points (H0, H2, H24, H48) on centrifuged plasma kept in contact with blood cells.The citrate showed the highest and most stable glucose values over time,with virtually no loss of glucose.Lower glucose values were observed at H0 and H2 with the NaF ($p < 0.05$).Transporting the heparin tube on ice combined with centrifugation at +4°C resulted in higher blood glucose values than the NaF at time points H0 and H2 ($p < 0.05$),although lower than the citrate($p < 0.05$);however, these values could not be maintained beyond 24 hours.Positive biases exceeding the desirable bias (1.8%) were observed between citrate and NaF tubes,as well as between heparinised and NaF tubes;this suggests a marked discrepancy and, consequently,a clinically significant difference between the various methods.For accurate blood glucose measurement,the citrate-NaF-EDTA tube is preferable in order to minimise glucose loss; however,a review of the diagnostic thresholds for diabetes mellitus must be carried out before its use.Cooling the sample is an effective alternative to compensate for the lack of citrate.

Introduction

Blood glucose is a key parameter for the diagnosis of diabetes mellitus, particularly in resource-limited settings. Accurate measurement of blood glucose is essential for appropriate patient management [1]. However, during the analytical process of glucose determination, several pre-analytical factors [2] may influence glucose concentration in the sample, including factors related to the patient's physical condition, those associated with sample handling and storage, and an intrinsic factor, namely glycolysis.

Glycolysis is a physiological process that occurs in the presence of blood cells and consists of glucose degradation through a series of enzyme-catalyzed chemical reactions, resulting in the production of energy and metabolites, including pyruvate, which is essential for the functioning of living organisms [3]. This process continues in vitro after blood collection and may be enhanced by prolonged contact time between blood cells and plasma [2,4,5], environmental temperature and hydrogen potential surrounding the blood cells [6], and sample cellularity [7,8].

When glycolysis is not controlled in blood samples, a decrease in plasma glucose concentration of 5–7% per hour may occur, corresponding to approximately 0.6 mmol/L or 0.1 g/L [9]. This may lead to misclassification of diabetic patients and, consequently, inappropriate clinical management.

To mitigate the impact of glycolysis during blood glucose analysis, various physical and chemical approaches have been recommended, including immediate separation of plasma from blood cells by centrifugation, the use of antiglycolytic additives, or immediate placement of collected blood samples on ice followed by centrifugation within 15 to 30 minutes [2,10,11]. The latter procedure has proven to be less practical in routine practice due to the required materials and equipment [12]. Among glycolysis inhibitors, sodium fluoride coated in blood collection tubes remains one of the most widely used in medical laboratories. Sodium fluoride (NaF) inhibits enolase, an enzyme involved in the final step of glycolysis [4]. However, several studies have demonstrated that the antiglycolytic effect of NaF becomes effective only two to four hours after phlebotomy [4,13,15], resulting in glucose loss in collected blood samples estimated at approximately 0.1 g/L per hour [16].

Consequently, recent recommendations from the American Diabetes Association (ADA) and the American Association for Clinical Chemistry (AACC), published in 2023 and updated in 2024, advocate the use of citrate for glycolysis inhibition as a replacement for NaF [1]. Citrate, through blood acidification, immediately halts in vitro glycolysis by inhibiting early glycolytic enzymes, including hexokinase and phosphofructokinase [14]. The antiglycolytic effect of citrate is maintained for approximately ten hours at room temperature [17]. However, the use of citrate in liquid form leads to dilution of the blood sample. The granular form of citrate, as contained in Vacuette® FC Mix tubes, overcomes this limitation. In addition to citrate, these tubes contain sodium fluoride associated with potassium oxalate (NaF/KOX) and ethylenediaminetetraacetic acid (EDTA) as an anticoagulant; the mixture of these additives is presented in the form of fine granules. According to several studies [11,15,18], in addition to ensuring complete inhibition of glycolysis through the combined inhibitory action of citrate and NaF, Vacuette® FC Mix tubes prevent the dilution effect [11].

Vacuette® FC Mix tubes are not available in the Democratic Republic of the Congo. To date, NaF/KOX-containing blood collection tubes remain the standard practice for blood glucose determination in the Biochemistry Laboratory of the University Clinics of Kinshasa.

This study aims to compare glucose stability among citrate–sodium fluoride–EDTA, lithium heparin, and sodium fluoride-coated blood collection tubes in a laboratory setting.

Materials and Methods

Study population

This prospective experimental study was conducted in accordance with the principles of the Declaration of Helsinki, revised in 2024, and was approved by the Ethical Committee of the School of Public Health at the University of Kinshasa (ESP/CE/42/2025).

A cohort of seventy participants, including both men and women, was consecutively recruited among voluntary blood donors attending a blood donation campaign organized in March 2025 by the Blood Bank of the University Clinics of Kinshasa, Democratic Republic of the Congo. The blood donors included in the study were those who were considered apparently healthy and non-diabetic according to routine pre-donation screening criteria [19] and provided informed consent.

Sample collection

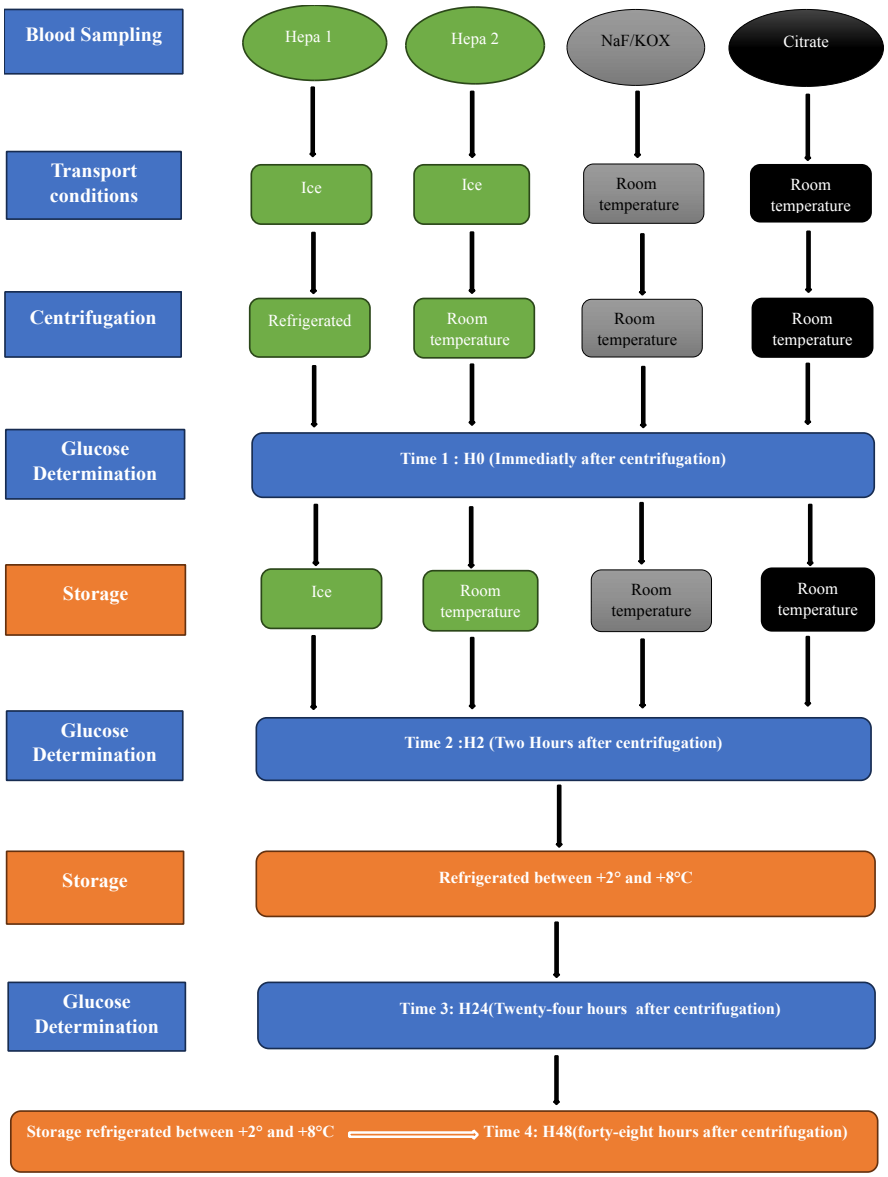
For each participant, after blood donation, peripheral venous blood was collected into four blood collection tubes: two lithium heparin–coated tubes (VACUUBLOOD®, provided by AMT Pharma, DR Congo), hereafter referred to as Hepa 1 and Hepa 2 tubes; one sodium fluoride and potassium oxalate–coated tube (VACUUBLOOD®, provided by AMT Pharma, DR Congo), referred to as the NaF/KOX tube; and one tube coated with a mixture of ethylenediaminetetraacetic acid (EDTA), sodium fluoride and potassium oxalate (NaF/KOX), and citric acid–sodium citrate (VACUETTE® FC Mix, provided by Greiner Bio-One, France), referred to as the citrate tube.

Sample transportation, pretreatment, and glucose analysis

From the sampling, the Hepa 1 tube was transported on ice to the Genetics Laboratory of the Faculty of Medicine at the University of Kinshasa within a maximum of 10 minutes for centrifugation at +4°C (MEGAFUGE 16R refrigerated centrifuge, Thermo Scientific, China). It was then sent to the Biochemistry Laboratory of the University Clinics of

Kinshasa within a maximum of 30 minutes for various blood glucose tests (15 minutes for centrifugation and 15 minutes for transportation).
Furthermore, the Hepa 2 tubes were transported on ice, while the Citrate and NaF tubes were transported at room temperature. These last three tubes were sent directly to the Biochemistry Laboratory of the University Clinics of Kinshasa for centrifugation at room temperature (SLOTEQ centrifuge, serial number W1920001906596, Finland) and for various blood glucose measurements within a maximum of 30 minutes. On centrifuged plasmas kept in contact with blood cells, Glucose measurements were performed using the MINDRAY BS240 analyzer (MINDRAY BS240, YX91002321, Shenzhen, China) with glucose oxidase–peroxidase reagents supplied by Mindray. Prior to glucose measurement, calibration and internal quality control procedures were performed using the Mindray Multi Sera Calibrator (lot 150124006) and Mindray Clinchem Multi Control level 1 (lot 0593224013) and level 2 (lot 059424010).Glucose concentrations were determined at four different time points: time zero (H0), immediately after centrifugation; time one (H2), 2 hours after centrifugation; time two (H24), 24 hours after centrifugation; and time three (H48), 48 hours after centrifugation.After H0 glucose measurement, the Hepa 1 tube was maintained in contact with ice, whereas Hepa 2, NaF/KOX, and citrate tubes were stored at room temperature. After H2 measurement, all primary tubes were stored at temperatures between +2 and +8 °C pending H24 and H48 glucose measurements.

Figure 1: The flowchart describe the pre analytical and the analytical process of glucose determination.



**Centrifugation: 2200 revolutions for 15 minutes (technical specifications of sampling tubes and compilation of data from similar studies).*

Data analysis

Statistical analyses were performed using Stata Statistical Software: Release 18 (StataCorp, 2023), and graphical representations were generated using R software, version 4.4.2 (2024). Comparisons between paired tubes were conducted using the Wilcoxon signed-rank test. Agreement between measurements was assessed using Bland–Altman plots, constructed for each tube pair comparison at each measurement time point. A Passing-Bablok regression was also performed to validate the results obtained using the Bland-Altmann plot. The effects of tube type and analysis time on blood glucose concentrations were evaluated using a generalized linear mixed model (GLMM). All analyses were interpreted using a 95% confidence interval, with the level of statistical significance set at 0.05.

Results

As blood glucose values were not normally distributed, they were described using the median and interquartile range (IQR). Table 1 presents the median blood glucose values obtained for the different tube types at the various measurement time points (H0, H2, H24, and H48).

Table 1: Distribution of blood glucose levels (Median and IQR) in mmol/L across different types of tube and at four time points.

Tubes	Median glucose (mmol/L)			
	H0 (IQR)	H2 (IQR)	H24 (IQR)	H48 (IQR)
Hepa1	5.47(1.13)	5.46(1.11)	5.20(1.28)	4.64(1.31)
Hepa2	5.38(1.17)	5.35(1.22)	4.61(1.27)	4.13(1.41)
NaF/KOX	5.32(1.21)	5.29(1.22)	5.26(1.27)	5.35(1.30)
Citrate	5.67(1.24)	5.60(1.18)	5.64(1.23)	5.63(1.30)

Table 2 and Table 3 illustrates respectively the statistical significance ($p < 0.05$) of the differences in blood glucose concentrations observed in citrate, Hepa 1, and Hepa 2 tubes compared with the NaF/KOX tube and the differences in blood glucose concentrations observed in NaF, Hepa 1, and Hepa 2 tubes compared with the Citrate tube.

Table 2: Median blood glucose(mmol/L) levels in different tube compared with the NaF tube (Wilcoxon test) at each time point.

Tubes	Median (mmol/L)	IQR	p_value
Hepa1			
H0	5.47	1.13	<0.001
H2	5.46	1.11	<0.001
H24	5.2	1.28	<0.001
H48	4.64	1.31	<0.001
Hepa2			
H0	5.38	1.17	<0.001
H2	5.35	1.22	0.049
H24	4.61	1.27	<0.001
H48	4.13	1.41	<0.001
Citrate			
H0	5.67	1.24	<0.001
H2	5.6	1.18	<0.001
H24	5.64	1.23	<0.001
H48	5.63	1.3	<0.001

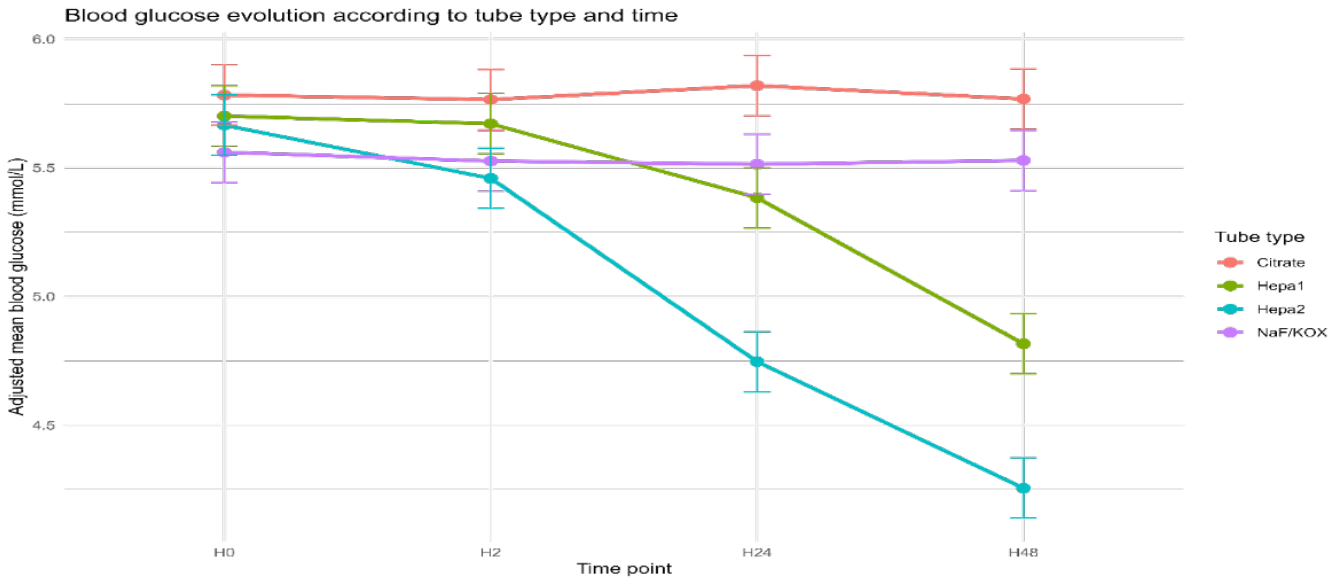
Table 3: Median blood glucose(mmol/L) levels in different tube compared with the Citrate tube (Wilcoxon test) at each time point.

Tubes	Median (mmol/L)	IQR	p_value
Hepa1			
H0	5.47	1.13	0.021
H2	5.46	1.11	0.001
H24	5.2	1.28	<0.001
H48	4.64	1.31	<0.001
Hepa2			
H0	5.38	1.17	<0.001
H2	5.35	1.22	<0.001
H24	4.62	1.27	<0.001
H48	4.14	1.41	<0.001
NaF			
H0	5.32	1.21	<0.001
H2	5.3	1.22	<0.001
H24	5.6	1.27	<0.001
H48	5.35	1.3	<0.001

The highest blood glucose concentrations were observed in the citrate tube at all measurement time points with a markedly significant difference(P<0.05) compared with the other tubes.
In the early measurement periods (H0 and H2), the Hepa 1 tube, for which the cold chain was maintained from collection to the H0 measurement, yielded blood glucose values lower than those of the citrate tube but higher than those of the NaF tube, a significant difference ($p < 0.05$) (Table 3).

The NaF tube gave low blood glucose values in the first two hours, with a significant difference compared to the citrate tube (P<0.05).
The Hepa1 and Hepa2 tubes showed decreasing blood glucose values beyond the H2 measurement, with a significant difference compared to the citrate tube (P<0.05).
Glucose stability over time across the different tube types is illustrated in Figure 2.

Figure 2: Adjusted mean blood glucose according to tube type and time of analysis, estimated using a generalized linear mixed model (GLMM).



Glucose stability was observed from H0 to H48 in the citrate and NaF tubes. In contrast, the Hepa 1 and Hepa 2 tubes showed a progressive and significant decrease in blood glucose ($p < 0.001$).
Figure 3 presents the kinetics of glucose degradation according to the analysis time and tube type, the initial glucose value measured at the first time point (H0) for each tube type.
Quantification of glucose losses for each tube at the different measurement time points, using H0 as the baseline, is presented in Table 4.

Figure 3: Loss of blood glucose over time for each tube type.

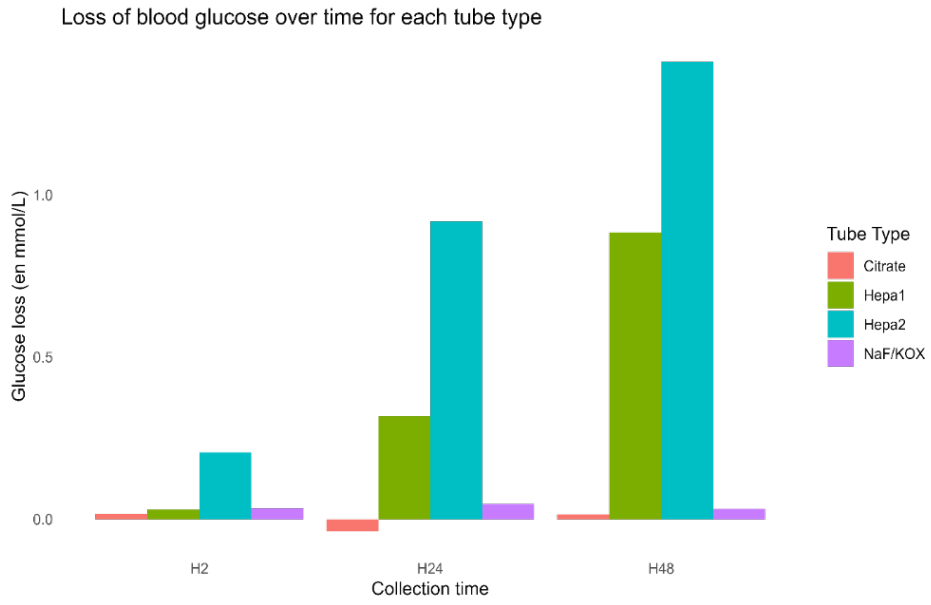


Table 4: Glucose loss comparison by tube and time. Estimated difference (%) compared to H0 with 95% CI.

	Estimate (%) [CI]	p_value	Signif
Hepa1			
H2	0.41 [-1.10; 1.92]	0.595	ns
H24	5.45 [3.93; 6.96]	<0.001	***
H48	15.62 [14.11; 17.13]	<0.001	***
Hepa2			
H2	3.64 [2.30; 4.99]	<0.001	***
H24	16.47 [15.12; 17.82]	<0.001	***
H48	25.26 [23.91; 26.61]	<0.001	***
NaF/Kox			
H2	0.62 [-0.18; 1.42]	0.1283	ns
H24	0.70 [-0.10; 1.49]	0.0872	ns
H48	0.52 [-0.28; 1.31]	0.2049	ns
Citrate			
H2	0.27 [-0.70; 1.24]	0.582	ns
H24	-0.75 [-1.71; 0.22]	0.132	ns
H48	0.22 [-0.75; 1.18]	0.662	ns

**** :P<0.05(Significative difference) ; ns : non significative difference.*

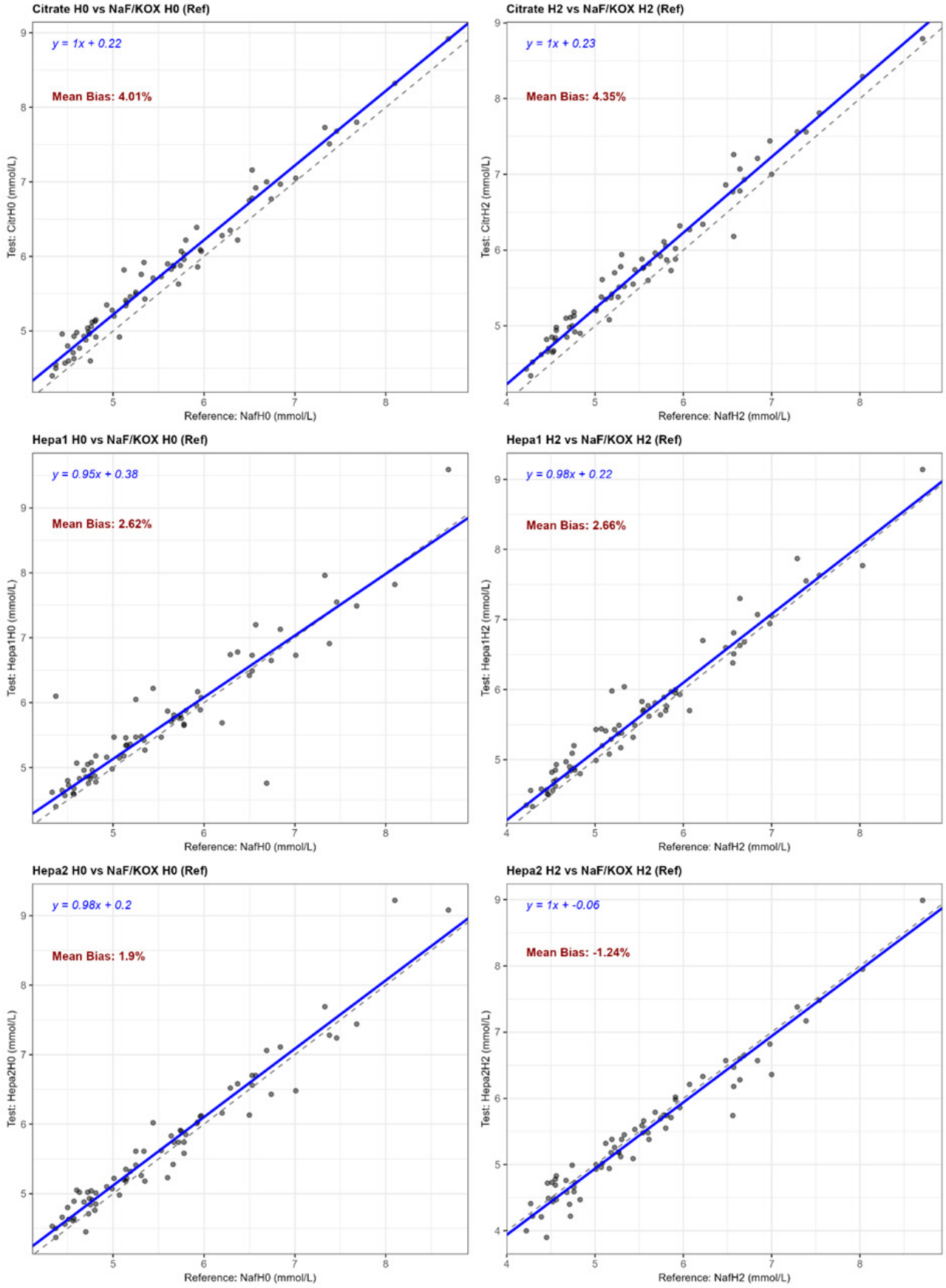
Over time, glucose losses were significant in the Hepa 1 and Hepa 2 tubes ($p < 0.05$); For the Hepa 1 tube, this was observed beyond H2, and for the Hepa 2 tube from the very start of the assay. In contrast, good glucose stability was observed with the Citrate and NaF tubes (losses ≤1%) and we noted a P value >0.05, non-significative.

Table 5: Agreement between alternative blood collection tubes and the NaF : A Bland-Altman analysis; Desirable bias at 1.8% [20].

Comparaison	Time	Bias Percent (%)	SD_Percent (%)	LoA_lower_percent	LoA_upper_percent
Citrate	H0	4.01	3.06	-1.99	10
Citrate	H2	4.35	3.02	-1.57	10.28
Hepa1	H0	2.62	7.01	-11.13	16.36
Hepa1	H2	2.03	3.75	-5.33	9.38
Hepa2	H0	1.9	3.85	-5.64	9.45
Hepa2	H2	-1.88	3.98	-9.67	5.92

Table 5 shows the degree of agreement between the citrate tubes and the NaF tubes, the Hepa 1 tubes and the NaF tubes, and the Hepa 2 tubes and the NaF tubes at the H0 and H2 measurement times. The comparison biases for each pair of tubes were compared with the target bias provided by the database of chemical performance specifications derived from the consensus of the European Federation of Clinical Chemistry and Laboratory Medicine and Westgard (target bias 1.8% for glucose) [20]. We note, in the first few hours, a discrepancy between citrate and NaF, as well as between heparinised tubes and NaF. The same conclusions are drawn using Passing-Bablok regression, as illustrated in the following figure. The statistical differences observed between the citrate-NaF and heparin-1-NaF tubes are likely, if not certainly, associated with actual clinical differences.

Figure 4: Passing-Bablok regression analysis and mean percentage bias of plasma glucose levels measured in different collection tubes compared to the NaF tube reference.



Discussion

In this study, the NaF tube was compared with the granular mixture of citrate, sodium fluoride, and ethylenediaminetetraacetic acid contained in the Vacuette® FC Mix tube (citrate tube), as well as with lithium heparin tubes stored on ice (Hepa 1 and Hepa2). The citrate tube yielded the highest blood glucose values compared with the other tube types, with near-zero and non-significant glucose losses over time. This demonstrates that blood acidification enables early, complete, and sustained inhibition of glycolysis for up to 48 hours. Indeed, each reagent contained in the citrate tube contributes to the prolonged antiglycolytic effect.

Furthermore, the citrate-containing tubes used in our study were in granular form. This is advantageous, as it avoids any dilution effects on glucose concentration compared with tubes containing citrate in liquid form [21,22]. This result is in line with previously published data Fokkert et al. [23], Uchida K et al [24], Gambino et al [25], Rebecca Carey et al. [26], Graziella Bonetti et al [27], Patil S et al [28], Gupta S et al [29] , Ridefelt P et al[30].

Higher glucose measures observed with the use of citrate coated tubes, raise some diabète mellitus diagnostic issues, depending on the population involved. The study by del Pino et al [31] reported a significant raise of positive gestational diabète mellitus after oral glucose tolerance test; but the difference was not significant in general patients . Riefelt P et al used data of patients from clinical routine health care and observed increased values of glycemia above the diabete decision limit after implementation of citrate coated tubes [30]. The overestimated plasma glucose level observed with citrate tube is not due to glucose increase, but only to the related more accurate glucose preservation in the blood sample. Consequently, usage of citrate tubes can increase the prevalence of impaired fasting glucose as well as the diagnosis of diabetes mellitus since glucose levels are closer to the diagnostic thresholds. In contrast, the study by Fisher M et al. [32] did not report on glucose overestimation with the citrate tube. The difference in methodology could explain the result size of the studied population and also samples storage at room temperature.

The same result was reported by Van der Hagen et al. [11], who included pregnant women referred for gestational diabetes screening. However, they observed that the blood glucose values obtained using the citrate tube correlated perfectly with those obtained using the NaF/EDTA tube, with the Bland-Altman plot showing a non-significant bias of 0.06 mmol/L before and after glucose loading during an oral glucose tolerance test in a cohort of pregnant women. In our study, we observed biases exceeding the acceptable limit when comparing the different tubes against NaF [20]. The assessment of citrate-NaF agreement revealed a bias of approximately 4% (Table 5), this bias having been analysed only for measurements taken at time points H0 and H2. Tiago Tizziani et al. in their study also observed that the two methods(Citrate and NaF/EDTA) Practically presented no correlation ($r = 0.15$)[33].In view of these results, citrate and NaF tubes do not agree, just as with the comparison of heparinised tubes and NaF. Whilst citrate tubes do provide optimal blood glucose values for better management of diabetes mellitus [1], there is a need to initiate discussions on adjusting the diagnostic thresholds for diabetes mellitus (DM) with these new (citrate) tubes.

Although the NaF tube showed slightly lower median blood glucose levels than the citrate tube, with a statistically significant difference ($P<0.001$), it demonstrated good glucose stability over time, matching the citrate tube, which showed optimal stability.

Given the delayed antiglycolytic effect of NaF, one would expect glucose losses to be significant over time, particularly at H2. However, Gambino et al. [25] reported significant glucose losses of 4.6% at H2 with the NaF tube. This discrepancy with our results may be explained by methodological differences. In Gambino et al., aliquots were stored at 37 °C, with the elevated temperature serving as a factor that exacerbates glycolysis.

During the first few hours, the Hepa 1 tube yielded higher median blood glucose values than the NaF tube ($P < 0.05$) and slightly lower values than the citrate tube ($P < 0.05$), with non-significant glucose losses ($P = 0.595 > 0.05$ at H2). This result may be explained by the fact that the Hepa 1 sample was transported in an ice slurry, centrifuged at +4 °C and kept in contact with ice until the H2 measurement, with the cold chain exerting an early inhibitory effect on the collected sample. However, this procedure proved less practical in routine practice due to the materials and equipment required, which are not always available in our setting [12]. Consequently, another laboratory had to be commissioned to carry out the refrigerated centrifugation of these samples from our cohort.

Recommendations advise separating plasma from cells within 15 to 30 minutes of blood collection [2]; however, strict adherence to this protocol is difficult to implement in routine practice. We encountered the same limitation in our study: although the plasma was kept in contact with the blood cells without decanting the sample, it took us a long time before the first measurement due to the distance to the centrifugation site. This could explain why the chilled Hepa1 tube, which was transported on crushed ice to the laboratory, proved less effective than the citrate tube.

Rebecca Carey et al. [26] reported similar glucose values between lithium heparin and sodium fluoride tubes, likely because both tubes were placed on ice immediately after collection, in contrast to our protocol.

This study is among the first conducted in our setting, as citrate mixture–coated tubes are recommended for blood glucose determination.

A limitation of the study is that the study population comprised only healthy individuals. Further studies of citrate tubes in our context, within a population suspected of having diabetes

mellitus or diagnosed with diabetes, would be necessary in order to better understand the clinical differences between citrate tubes and NaF tubes and to draw definitive conclusions. Cases of diabetes mellitus would likely increase with citrate tubes, given their anti-glycolytic efficacy in vitro.

Conclusion

In conclusion, among NaF and heparin containing tubes, the citrate-NaF-EDTA coated tubes appear to have the most stable glycolyse inhibitor effect over time, in our laboratory-based study. If citrate tubes are not available, using an ice slurry to transport samples, or better still, refrigerated centrifugation, would be a good alternative. The current diagnostic thresholds based on blood samples collected in tubes containing only NaF are probably too low.A reassessment of the decisive diagnostic thresholds will have to be carried out with the citrate tube and, why not, also with the tube preserved in ice.

Funding

The authors received no financial support for the research, authorship, and publication of this article.

Data Availability

All data are available from the corresponding author upon reasonable request.

Declaration of conflict of Interest

Authors do not have any conflict of interest.

Ethical Approval

Ethical approval was obtained by the research panel (ethics number ESP/CE/42/2025).

Authors contributions

-Ongona Sheko contributed to conceptualization, data curation, formal analysis, investigation, validation, methodology, visualization, writing – original draft.
-Ngole Mamy contributed to conceptualization, data curation, formal analysis, investigation, validation, methodology, visualization, writing – original draft.
- Pambu Dahlia and Tshibuela Beya Dophie contributed to review and editing.
-Mukadi Patrick and Beya Michael participated in the statistical analysis.

Acknowledgement

We acknowledged the efforts of all laboratory staff during our research study.

References

1. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Lernmark Å, Metzger BE, Nathan DM, Kirkman MS. Guidelines, and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Diabetes Care. 2023;46(10):e151–e199. doi:10.2337/dci23-0036. PMID: 37471273; PMCID: PMC10516260.
2. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Kirkman MS, Lernmark A, Metzger BE, Nathan DM. Guidelines, and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Clin Chem. 2011;57(6):e1–e47. doi: 10.1373/clinchem.2010.161596. PMID: 21617152.
3. Chaudhry R, Varacallo MA. Biochemistry: Glycolysis. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan–. Updated 8 Aug 2023. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482303/
4. Mikesh LM, Bruns DE. Stabilization of glucose in blood specimens: mechanism of delay in fluoride inhibition of glycolysis. Clin Chem. 2008;54(5):930–932. doi:10.1373/clinchem.2007.102160. Erratum in: Clin Chem. 2008;54(7):1261. PMID: 18443184.
5. Daimane D, Basu S. Efficacy, and clinical utility of sodium fluoride tubes versus serum tubes for blood glucose estimation. Cureus. 2024;16(10):e72641. doi:10.7759/cureus.72641. PMID: 39610588; PMCID: PMC11604175.
6. Study Smarter GmbH. Factors Affecting Enzyme Activity. 2024. Available from: https://www.studysmarter.fr
7. Ybarra J, Isern J. Leukocytosis-induced artifactual hypoglycemia. Endocr J. 2003;50(4):481–482. doi:10.1507/endocrj.50.481. PMID: 14599125.
8. Bossard V, Sauvageon Y, Fraissinet F, Dreyfus B, Thuillier R, Hauet T. Fausse hypoglycémie induite par les paraprotéines lors de la mesure du glucose par la méthode de l'hexokinase. Ann Biol Clin. 2019;77(4):439–445. doi:10.1684/abc.2019.1465.
9. Stahl M, Jørgensen LG, Hyltoft Petersen P, Brandslund I, de Fine Olivarius N, Borch-Johnsen K. Optimization of preanalytical conditions and analysis of plasma glucose. 1. Impact of the new WHO and ADA recommendations on diagnosis of diabetes mellitus. Scand J Clin Lab Invest. 2001;61(3):169–179. doi:10.1080/003655101300133612. PMID: 11386604.
10. World Health Organization, International Diabetes Federation. Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia. Geneva: WHO; 2006. ISBN: 9241594934.
11. van der Hagen EAE, Fokkert MJ, Kleefman AMD, Thelen MHM, van den Berg SAA, Slingerland RJ. Technical and clinical validation of the Greiner FC-Mix glycaemia tube. Clin Chem Lab Med. 2017;55(10):1530–1536. doi:10.1515/cclm-2016-0944. PMID: 28284032.
12. van den Berg SA, Thelen MH, Salden LP, van Thiel SW, Boonen KJ. It takes acid, rather than ice, to freeze glucose. Sci Rep. 2015;5:8875. doi:10.1038/srep08875. PMID: 25748167; PMCID: PMC4352852.
13. Chan AY, Cockram CS, Swaminathan R. Effect of

delay in separating plasma for glucose measurement upon the interpretation of oral glucose tolerance tests. Ann Clin Biochem. 1990;27(Pt 1):73–74. doi:10.1177/000456329002700115. PMID: 2310160.
14. Gambino R. Sodium fluoride: an ineffective inhibitor of glycolysis. Ann Clin Biochem. 2013;50(Pt 1):3–5. doi:10.1258/acb.2012.012135. PMID: 23129725.
15. Stapleton M, Daly N, O'Kelly R, Turner MJ. Time and temperature affect glycolysis in blood samples regardless of fluoride-based preservatives: a potential underestimation of diabetes. Ann Clin Biochem. 2017;54(6):671–676. doi:10.1177/0004563216682978. PMID: 28084093.
16. Chan AY, Swaminathan R, Cockram CS. Effectiveness of sodium fluoride as a preservative of glucose in blood. Clin Chem. 1989;35(2):315–317. PMID: 2914384.
17. Waring WS, Evans LE, Kirkpatrick CT. Glycolysis inhibitors negatively bias blood glucose measurements: potential impact on the reported prevalence of diabetes mellitus. J Clin Pathol. 2007;60(7):820–823. doi:10.1136/jcp.2006.039925. PMID: 17596547; PMCID: PMC1995804.
18. Griebenow S, de Kozlowski S. Prevention of glycolysis with the VACUETTE FC Mix tube—a precise diagnosis of diabetes. Diabetes. 2024;73(Suppl 1):726–P. doi:10.2337/db24-726-P.
19. Arrêté du 17 décembre 2019 fixant les critères de sélection des donneurs de sang. Updated 16 Mar 2022. NOR: SSAP1936572A. Available from: https://www.legifrance.gouv.fr/loda/id/JORFTEXT000039667225/
20. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, Minchinela J, Perich C, Simon M. "Current databases on biologic variation: pros, cons and progress." Scand J Clin Lab Invest 1999;59:491-500.
21. van der Hagen EA, Kleefman AM, Thelen MH, van den Berg SA. Normalisation issues in glucose measurements using phlebotomy tubes with liquid additives. Clin Chem Lab Med. 2017;55(1):e1–e3. doi:10.1515/cclm-2016-0225. PMID: 27362962.
22. Juricic G, Saracevic A, Kopcinovic LM, Bakliza A, Simundic AM. Evidence for clinically significant bias in plasma glucose between liquid and lyophilized citrate buffer additives. Clin Biochem. 2016;49(18):1402–1405. doi:10.1016/j.clinbiochem.2016.03.006. PMID: 27040902.
23. Fokkert M. Glucose (monitoring): from bench to real world experiences [Thesis]. University of Groningen; 2022. doi:10.33612/diss.207217778.
24. Uchida K, Matuse R, Toyoda E, Okuda S, Tomita S. A new method of inhibiting glycolysis in blood samples. Clin Chim Acta. 1988;172(1):101–108. doi:10.1016/0009-8981(88)90125-8. PMID: 2966020
25. Gambino R, Piscitelli J, Ackattupathil TA, Theriault JL, Andrin RD, Sanfilippo ML, Etienne M. Acidification of blood is superior to sodium fluoride alone as an inhibitor of glycolysis. Clin Chem. 2009;55(5):1019–1021. doi:10.1373/clinchem.2008.121707. PMID: 19282354.
26. Carey R, Lunt H, Heenan HF, Frampton CM, Florkowski CM. Collection tubes containing citrate stabilizer overestimate plasma glucose compared with samples undergoing immediate plasma separation. Clin Biochem. 2016;49(18):1406–1411. doi:10.1016/j.clinbiochem.2016.05.017. PMID: 27234599.
27. Bonetti G, Carta M. The new Greiner FC-Mix tubes equal the old Terumo ones and are useful as glucose stabilizers after prolonged storage of samples. Biochem Med (Zagreb). 2017;27(3):030901. doi:10.11613/BM.2017.030901. PMID: 29018304; PMCID: PMC5628000.
28. Patil S, Reshma DC. Determining the efficacy of citrate buffer tubes compared to sodium fluoride tubes as inhibitors of glycolysis in glucose estimation. Int J Clin Biochem Res. 2019;6(3):299–302. doi:10.18231/j.ijcbr.2019.066.
29. Gupta S, Gupta AK, Verma M, Singh K, Kaur A, Chopra B, Kaur V. Comparison of plasma glucose levels obtained in sodium fluoride and citrate buffer tubes at a tertiary care hospital in Punjab. Int J Appl Basic Med Res. 2016;6(1):50–53. doi:10.4103/2229-516X.174010. PMID: 26958523; PMCID: PMC4765275.
30. Ridefelt P, Åkerfeldt T, Helmersson-Karlqvist J. Increased plasma glucose levels after change of recommendation from NaF to citrate blood collection tubes. Clin Biochem. 2014;47(7–8):625–628. doi:10.1016/j.clinbiochem.2014.02.022. PMID: 24631175.
31. Del Pino IG, Constanso I, Mourín LV, Safont CB, Vázquez PR. Citric/citrate buffer: an effective antiglycolytic agent. Clin Chem Lab Med. 2013;51(10):1943–1949. doi:10.1515/cclm-2012-0735. PMID: 23612543.
32. Fischer MM, Hannemann A, Winter T, Schäfer C, Petersmann A, Nauck M. Relative efficacy of different strategies for inhibition of in vitro glycolysis. Clin Chem. 2021;67(7):1032–1034. doi:10.1093/clinchem/hvab071. PMID: 34104943.
33. Tiago Tizziani, Diogo Alexandre Siebert, Tassiana Pancotte, Caio Mauricio Mendes de Cordova Department of Pharmaceutical Sciences, University of Blumenau, Blumenau, Brazil. Effect of Addition of Sodium Citrate in Plasma Extraction for Venous Blood Glucose Determination. International Journal of Clinical Medicine > Vol.5 No.11, May 2014 DOI: 10.4236/ijcm.2014.511085

Supplemental material

*Lithium heparin tubes: VACUUBLOOD®, provided by AMT Pharma, DR Congo, *Fluoride and potassium oxalate tube: VACUUBLOOD®, provided by AMT Pharma, DR Congo. *Tube coated with a mixture of ethylenediaminetetraacetic acid (EDTA), sodium fluoride and potassium oxalate (NaF/KOX), and citric acid–sodium citrate: VACUETTE® FC Mix, provided by Greiner Bio-One, France-

* Centrifuges: -MEGAFUGE 16R refrigerated centrifuge, Thermo Scientific, China. -SLOTEQ centrifuge, serial number W1920001906596, Finland.
*The MINDRAY BS240 analyzer: MINDRAY BS240, YX91002321, Shenzhen, China with glucose oxidase–peroxidase reagents supplied by Mindray.

*Calibration and internal quality control: Mindray Multi Sera Calibrator (lot 150124006), from china and Mindray Clinchem Multi Control level 1 (lot 0593224013) and level 2 (lot 059424010) from China.

Research Article

Statistical Analysis of the Performance of Open-Access Language Models: A Tool in the Clinical Laboratory

Micaela Avellaneda[1], Pamela De Francesco[2], Nilda Fink[3*]

[1]*B.Sc. in Biochemistry Federacion Bioquimica de la Provincia de Buenos Aires*
[2]*Library scientist Fundación Bioquímica Argentina (FBA)*
[3]*PhD, MSc in Biochemistry and Pharmacy. Fundación Bioquímica Argentina (FBA)*

Article Info

[*]***Corresponding Author:***
Nilda Fink
Fundación Bioquímica Argentina (FBA) PROES La Plata, Buenos Aires, Argentina
E-mail: nfink@fba.org.ar

Keywords

Artificial intelligence, language models, comparative evaluation, reproducibility, clinical laboratory

Abstract

Introduction: Artificial intelligence (AI) has been progressively incorporated into the clinical analysis laboratory, optimizing data management, detection of complex patterns, and image-based diagnostic support. Among emerging tools, Large Language Models (LLMs) represent a significant advance by enabling natural language processing for clinical, analytical, and academic tasks. However, their performance is heterogeneous and depends on model architecture, training, and contextual fine-tuning. Objective: To compare the performance of eight openly accessible large language models (LLMs) applied to the biomedical domain by assessing their output quality and inter-rater consistency.

Materials and Methods: A cross-sectional comparative study was carried out. Eight LLMs (ChatGPT-4, DeepSeek V3.1, Gemini, Copilot GPT-5, Claude 3.5 Haiku, Perplexity AI, DeepAI, and Poe) were queried using six standardized questions regarding the analytical evolution of blood glucose measurement. Responses were evaluated by two independent automated systems (Text Cortex AI and Grok Code Fast 1) using a Likert scale from 1 to 10 and considering eight criteria of textual and scientific quality. Friedman and Wilcoxon–Holm tests were used for global and pairwise comparisons, and Pearson and Spearman coefficients were calculated to assess inter-rater agreement.

Results: Significant differences were identified among models ($p < 0.001$). The LLMs Gemini, DeepSeek, and ChatGPT-4 obtained the highest scores (medians > 8.5), whereas DeepAI and Poe showed lower performance. Inter-rater correlation based on median scores showed a markedly strong linear association ($r = 0.91$; $p = 0.0019$) and a robust ordinal agreement ($\rho = 0.83$; $p = 0.011$). The least-squares regression line ($E2 = 1.37 \cdot E1 – 3.25$) indicated a sustained positive relationship between the two systems, with balanced dispersion around the linear model. These findings confirm the stability of the evaluation pattern and support the reproducibility of the method.

Conclusion: The LLMs with greater complexity and contextual fine-tuning demonstrated superior accuracy, coherence, and reproducibility; however, it is suggested that LLMs may serve as complementary tools depending on the clinical or research context. Their use in healthcare requires methodological validation and expert oversight to ensure safety and clinical applicability within established ethical and regulatory frameworks.

Introduction

Artificial intelligence (AI) has become a key pillar in the clinical laboratory due to its ability to analyze large volumes of data, identify complex patterns, and support both clinical and operational decision-making [1,2]. Its applicability extends across all three phases of the diagnostic process. In the pre-analytical phase, recognized as the most vulnerable stage of the diagnostic process, AI and data analytics help to reduce frequent errors (such as those related to patient identification or sample handling), optimize simple traceability, and promote the standardization of procedures [3-5]. In the analytical phase, machine learning approaches have been applied to quality control, outlier detection, and automated validation, thereby improving the reliability and consistency of reported results [2,6]. In the post- analytical phase, integration with EHR systems enables risk stratification and clinical decision support; for instance, predictive models for sepsis can anticipate septic shock with robust performance (AUC ≈ 0.83) prior to clinical onset [7]. Additionally, the use of indirect data-driven methods to estimate reference intervals has expanded. A recent study in a Puerto Rican population applied unsupervised learning and demonstrated age- and sex-specific variations [8]. Outside the strictly analytical setting, AI-based triage systems in emergency departments show potential for improving prioritization and process consistency, however, they require multicenter validation and standardized reporting [9,10]. Finally, large language models (LLMs) provide a natural-language interface that is useful for documentation, synthesis, and diagnostic support. Recent literature and systematic reviews describe their applications, limitations, and safety considerations as well as ongoing clinical trials evaluating their implementation [11-13]. In this context, it is crucial to systematically and critically assess the performance of these tools in tasks that require historical knowledge, factual accuracy, and clarity of exposition. This study aims to compare the performance of eight language models by analyzing their responses to six standardized questions addressing the historical evolution and methodological development of glucose measurement. The evaluation includes an assessment of historical accuracy, the reliability of verifiable sources, and the precision, clarity, and comprehensiveness of the generated content. It also examines originality, factual coherence, argumentative depth, and adherence to academic style based on predefined criteria. Inter-rater agreement will be quantified using correlation coefficients to ensure methodological rigor. In summary, this approach intends to provide a robust estimation of the utility and limitations of large language models as supportive tools in clinical, educational, and research settings.

Materials and Methods

Study Design

A cross-sectional comparative study with non-parametric statistical analysis was conducted to evaluate the performance of eight LLMs.

Analyzed Models

The models included in the analysis were selected because they represent the main architectures and technological families of freely accessible LLMs available at the time of the study:
M1: ChatGPT-4 (OpenAI)
M2: DeepSeek V3.1
M3: Gemini (Google DeepMind)
M4: Copilot Smart GPT-5 (Microsoft)
M5: Claude 3.5 Haiku (Anthropic)
M6: Perplexity AI
M7: DeepAI
M8: Poe (Quora)
Each model was queried on October 10, 2025, under standardized conditions using a set of six sequential questions designed to assess knowledge of the historical evolution, analytical foundations, and technological applications of blood glucose measurement. These questions addressed, among other aspects, the initial description of the colorimetric method, the introduction of the enzymatic method, key milestones in laboratory automation, and recent innovations in portable devices and integrations with artificial intelligence.
The responses obtained were compiled in text format and evaluated automatically and independently by two AI-based systems: E1: Text Cortex IA and E2: Grok Code Fast 1. Both are freely accessible tools that function as independent evaluators, assigning scores on a Likert scale from 1 to 10 (with 10 being the highest rating) based on eight predefined criteria: historical accuracy, exhaustiveness, narrative clarity, verifiable sources, originality, factual coherence, argumentative depth, and academic style.

Statistical Analysis

Non-parametric methods were applied, as they are suitable for ordinal data and asymmetric distributions. Descriptive measures included the median and interquartile range (IQR). Overall comparisons of model scores were conducted using the Friedman test. Post-hoc analyses were performed with the paired Wilcoxon test, applying Holm's correction to control for multiple-comparison error.
Inter-rater consistency was assessed using Pearson (r) and Spearman (ρ) correlation coefficients. The significance level was set at $\alpha = 0.05$.
All analyses were performed using SPSS Statistics version 20.0.

Results

The comparative analysis demonstrated statistically significant differences among the evaluated models for both scoring systems (E1 and E2), with Friedman test p-values < 0.001. This supports the presence of true variability in the models' performance when addressing standardized blood glucose measurement queries.

According to the scores assigned by evaluator E1, models M3, M2, and M1 constituted the highest-performing group (Table 1). Their medians and interquartile ranges (IQR: Q3–Q1) were 8.9 (0.9), 8.7 (1.0), and 8.6 (1.0), respectively, reflecting a high degree of narrative consistency, argumentative clarity, and historical accuracy (Figure 1). Models M4 and M5 showed intermediate performance, whereas M6, M7, and M8 ranked in the lower tier, with scores below 7 points on the Likert scale. The post-hoc Wilcoxon analysis with Holm correction confirmed significant differences between the top-performing group (M3–M2–M1) and the lower-performing group (M7–M8–M6), validating the robustness of the resulting ranking.

Evaluator E2 identified a similar pattern, although with slight variations in internal ordering (Table 1). In this case, M1 achieved the highest median—9.4 (0.7)—followed by M2 with 8.6 (0.7) and M3 with 8.0 (1.4) (Figure 2). Models M4 and M5 maintained acceptable values, whereas M6, M7, and M8 exhibited significantly lower performance ($p < 0.05$ in pairwise comparisons M1 > M8; M2 > M7; M2 > M8).

Inter-rater correlation calculated from median scores demonstrated a very high linear association ($r = 0.91$; $p = 0.0019$) and a robust ordinal agreement ($\rho = 0.83$; $p = 0.011$). The least-squares regression line ($E2 = 1.37 \cdot E1 - 3.25$) indicated a sustained positive relationship between the two systems, with balanced dispersion around the linear model (Figure 3). These results confirm the stability of the evaluation pattern and support the reproducibility of the method.

The joint analysis of the assessments provided by both evaluators indicates that models M1, M2, and M3 constitute the group with the highest overall performance. This group is characterized by producing coherent and well-structured responses, demonstrating strong contextual command of scientific information, maintaining high levels of narrative clarity, exhibiting low dispersion (reflected in a narrow interquartile range) and showing a high degree of consistency across independent evaluators.

In contrast, models with simpler architectures (M6–M8) exhibited limitations in factual precision, use of verifiable references, and argumentative development—critical aspects in biomedical applications (Table 3).

Table 1: Comparison of Model Performance: Text Cortex vs. Grok Code Fast 1.

Model	Median [IQR] Text Cortex *	Interpretation (Text Cortex)	Median [IQR] Grok Code Fast 1 **	Interpretation (Grok Code Fast 1)
M1	8.6 [1.0]	Robust and well-balanced performance	9.4 [0.7]	Highest overall performance
M2	8.7 [1.0]	Stable coherence and depth	8.6 [0.7]	High coherence
M3	8.9 [0.9]	Maximum narrative clarity	8.0 [1.4]	Slight variability
M4	8.1 [0.9]	Good intermediate performance	8.1 [0.9]	Good stability
M5	7.3 [1.2]	Moderate variability	6.9 [1.3]	Moderate dispersion
M6	6.8 [1.2]	Lower but acceptable performance	6.9 [0.9]	Intermediate performance
M7	6.6 [0.7]	Limited accuracy	5.6 [1.2]	Low accuracy
M8	6.4 [0.7]	Low and homogeneous performance	4.8 [1.4]	Significantly lower performance
Results	* Text Cortex evaluator: χ^2 (7) = 46.97; $p < 0.001$ → significant differences. Post-hoc (Wilcoxon–Holm): significant differences between the top-performing group (M3–M2–M1) and the lower group (M7–M8–M6).		** Grok Code Fast 1 evaluator: χ^2 (7) = 53.88; $p < 0.001$ → highly significant differences. Post-hoc (Wilcoxon–Holm): M1 > M8 ($p = 0.0413$); M2 > M7 ($p = 0.0210$); M2 > M8 ($p = 0.0108$).	

Values are expressed as median [interquartile range, IQR] on a Likert scale of 1 to 10, where 10 represents the highest possible score (maximum clarity, coherence, and analytical precision).

The Friedman and Wilcoxon–Holm tests were applied to identify global and pairwise differences among models, respectively.

Table 2: Inter-Rater Correlation Between E1 and E2.

Parameter	Value	Interpretation
Pearson's r	0.91	Very high linear correlation (quantitative consistency)
Spearman's ρ	0.83	Robust ordinal correlation (agreement in ranking order)

Table 3: Comparative Summary of Overall Performance by Evaluator.

Evaluator	Overall Median	Overall IQR	Significance (Friedman)
E1	7.7	[6.6–8.8]	$p < 0.001$
E2	7.6	[6.3–8.7]	$p < 0.001$

Figure 1: Comparative Performance Ranking.

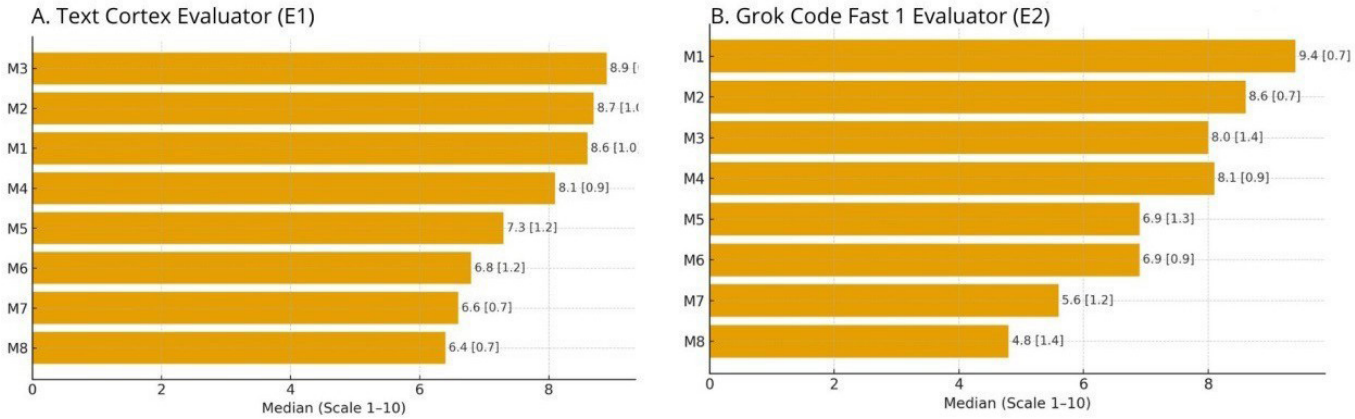
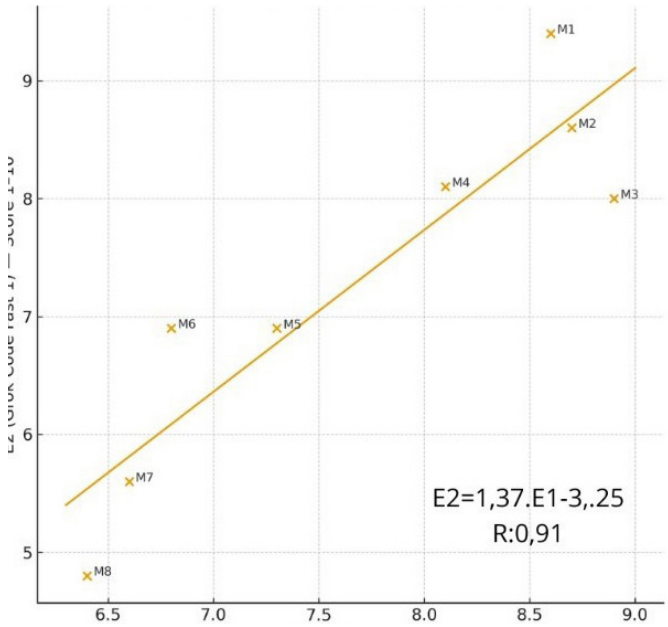


Figure 2: Inter-Rate-Correlation. Text Cortex vs Grok Code Fast 1.



Discussion

The results of this study demonstrate that the evaluated LLMs - particularly models M1, M2, and M3 - exhibit a clear advantage in both analytical capacity and narrative generation compared with general-purpose or minimally trained models. This superiority may be attributed to three key factors: (a) the multimodal scale of training (text and possibly additional modalities), (b) biomedical contextual calibration, and (c) optimization of semantic understanding. Overall, these components enhance stability and accuracy in clinical reasoning tasks, giving specialized models a more robust capacity to process complex information and generate coherent responses within the biomedical domain.

The use of non-parametric tests - such as the Friedman and Wilcoxon–Holm tests - was particularly appropriate for comparing non-normal distributions and minimizing the influence of outliers. Reporting medians and interquartile ranges (IQR) provided a bias-resistant characterization of central tendency, offering an accurate representation under conditions of high variability. Additionally, the incorporation of two independent automated evaluators strengthened the objectivity of the analysis and reduced the risk of subjective bias, demonstrating that automated systems can faithfully reproduce scientific evaluation criteria.

The high inter-rater correlation observed in this study confirms the reproducibility and consistency of the evaluation process. The observed linear concordance ($r = 0.91$; $p = 0.0019$) indicates that both systems largely agree on the magnitude of assigned scores, while the ordinal correlation ($\rho = 0.83$; $p = 0.011$) demostrates strong agreement in the relative ranking of models. These values fall within the range considered strong evidence of validity and stability in reproducibility studies, as reported in methodological literature [14].

The adjusted regression line ($E2 = 1.37 \cdot E1 – 3.25$) exhibited a pronounced positive slope, suggesting that variations in scores assigned by Text Cortex are proportionally reflected in the scores assigned by Grok Code Fast 1. The balanced distribution of points around the fitted line confirms the absence of systematic evaluator-specific biases and indicates that the observed dispersion arises from inherent variability in LLM performance. These results reinforce the methodological rigor of the study, demonstrating that evaluation was consistent, stable, and free of systematic scoring effects.

Our findings are consistent with recent evidence showing that LLM performance depends significantly on model architecture, training volume, and the degree of fine-tuning to the biomedical domain. For instance, one review study reported that LLMs can answer free-text medical queries with promising levels of aptitude but still show heterogeneous results [15]. In another study, which evaluated an LLM on the United States Medical Licensing Examination a performance exceeding 85% accuracy was reported, highlighting the model’s capacity for contextualized clinical reasoning [16]. Similarly, recent methodological reviews indicate that factual accuracy and narrative consistency in advanced models depend not only on training corpus size but also on hallucination control and reinforcement with clinically verified data via Reinforcement Learning with Human Feedback [17]. These factors may explain the observed superiority models M1 and M3 over less specialized models such as M6 and M8, which lack systematic biomedical fine-tuning.

The integration of LLMs with structured knowledge bases (e.g., UMLS, SNOMED CT, PubMedQA) has been shown to improve factual coherence and reduce interpretive errors, reinforcing the contextual validity of their responses [15]. This observation aligns with the results of the present study: the models with greater semantic integration and medical contextualization - M1 and M3 - achieved the highest clarity and accuracy scores.

From a statistical standpoint, the choice of non-parametric methods is well supported in the AI evaluation literature, given the frequent non-normality of score distributions between models [18]. Likewise, the high inter-rater correlation obtained is consistent with values reported in previous studies as markers of reproducibility in automated evaluation contexts [14].

From a broader perspective, our results reinforce the premise that the quality of AI in healthcare depends not only on algorithmic scale but also on ethical, regulatory, and human supervision frameworks. In this regard, the findings align with conclusions established in other works [15,19]. Consequently, the integration of AI into clinical laboratory workflows requires prospective validation, documentation traceability, and expert review before implementation in critical medical decisions.

From an applied perspective, the results confirm that LLMs can play a complementary role in the generation, analysis, and interpretation of biomedical knowledge, provided their performance is validated through objective, reproducible, and comparative criteria. In particular, model M1 emerges as the most consistent in precision and conceptual clarity, suggesting potential applications in clinical laboratory support, scientific writing, and medical education.

Nevertheless, their use should be conceived as a form of cognitive assistance intended to complement, not to replace professional human judgment and critical review. In high-stakes settings, the presence of biases, inaccuracies, or errors may carry serious consequences.

Among the main limitations of this study, it should be noted that the analyzed responses represent a single temporal snapshot of performance; therefore, potential variations due to model updates or architectural modifications were not evaluated. Future research should incorporate longitudinal analyses to examine performance evolution across new versions or training adjustments. Additionally, we recommend including further metrics of semantic coherence, expert cross-validation, and thematic sensitivity testing in specific biomedical areas such as oncology, genetics, or molecular diagnostics, in order to strengthen the evaluative framework

and consolidate clinical applicability.

Conclusions

The results of this study indicate that M1 achieved the highest overall performance, followed by M2 and M3. However, it is suggested that LLMs may operate in a complementary manner, as one model may outperform another in specific criteria, providing differentiated strengths depending on the evaluation conducted. The correlation analysis revealed a high level of consistency between the two automated evaluators ($r = 0.91$; $\rho = 0.83$), reinforcing the methodological robustness of the study. The fitted regression line, $E2 = 1.37 \cdot E1 - 3.25$, showed a stable positive relationship between the systems, with no evidence of differential bias. Taken together, these findings confirm that the inter-model comparisons were carried out under reliable and reproducible evaluation conditions.

These results suggest that LLMs represent promising tools for analytical and documentary support in the clinical laboratory context, provided their use is accompanied by appropriate ethical, regulatory, and methodological safeguards.

Ethical approval

This study did not involve human or animal subjects and therefore did not require ethical approval, in accordance with the Declaration of Helsinki (different versions).

Informed consent

Not applicable.

Conflict of interest

None declared.

Funding

This study did not receive financial support from public or private agencies.

Use of AI

This manuscript was drafted with the assistance of ChatGPT GPT-5 (OpenAI, 2025) for language refinement. The authors verified the accuracy and coherence of all content.

Author Contributions (CRediT)

Role	Author(s)
Conceptualization	Nilda Fink, Micaela Avellaneda, Pamela De Francesco
Methodology	Nilda Fink, Micaela Avellaneda, Pamela De Francesco
Formal analysis	Nilda Fink, Micaela Avellaneda, Pamela De Francesco
Investigation	Nilda Fink, Micaela Avellaneda, Pamela De Francesco
Validation	Nilda Fink
Writing – original draft	Micaela Avellaneda, Pamela De Francesco
Writing – review & editing	Nilda Fink, Micaela Avellaneda, Pamela De Francesco
Supervision	Nilda Fink

Data Availability and Transparency Statement

All data supporting the findings of this study are available upon reasonable request to the corresponding author. Anonymized datasets, scoring matrices, and figure source files are preserved for verification and reproducibility. The authors affirm full transparency in data handling, statistical analysis, and manuscript preparation, in accordance with IFCC and COPE guidelines.

References

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi: 10.1038/s41591-018-0300-7.
2. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med. 2019;25(1):24-29. doi: 10.1038/s41591-018-0316-z.
3. Van Moll C, Egberts T, Wagner C, Zwaan L, Ten Berg M. The Nature, Causes, and Clinical Impact of Errors in the Clinical Laboratory Testing Process Leading to Diagnostic Error: A Voluntary Incident Report Analysis. J Patient Saf. 2023;19(8):573-579. doi: 10.1097/PTS.0000000000001166.
4. Nordin N, Ab Rahim SN, Wan Omar WFA, Zulkarnain S, Sinha S, Kumar S, Haque M. Preanalytical Errors in Clinical Laboratory Testing at a Glance: Source and Control Measures. Cureus. 2024;16(3):e57243. doi: 10.7759/cureus.57243.
5. Simundic AM, Lippi G. Preanalytical phase--a continuous challenge for laboratory professionals. Biochem Med (Zagreb). 2012;22(2):145-149. doi: 10.11613/bm.2012.017.
6. Xie H, Jia Y, Liu S. Integration of artificial intelligence in clinical laboratory medicine: Advancements and challenges. Interdisciplinary Medicine 2024;2. doi: 10.1002/inmd.20230056.
7. Henry KE, Hager DN, Pronovost PJ, Saria S. A targeted real-time early warning score (TREWScore) for septic shock. Sci Transl Med. 2015;7(299):299ra122. doi: 10.1126/scitranslmed.aab3719.
8. Velev J, LeBien J, Roche-Lima A. Unsupervised machine learning method for indirect estimation of reference intervals for chronic kidney disease in the Puerto Rican population. Sci Rep. 2023;13:17198. Doi: 10.1038/s41598-023-43830-3.

9. Tyler S, Olis M, Aust N, Patel L, Simon L, Triantafyllidis C, et al. Use of Artificial Intelligence in Triage in Hospital Emergency Departments: A Scoping Review. Cureus. 2024;16(5):e59906. doi: 10.7759/cureus.59906.
10. Porto BM. Improving triage performance in emergency departments using machine learning and natural language processing: a systematic review. BMC Emerg Med. 2024;24(1):219. doi: 10.1186/s12873-024-01135-2.
11. Meng X, Yan X, Zhang K, Liu D, Cui X, Yang Y, et al. The application of large language models in medicine: A scoping review. iScience. 2024;27(5):109713. doi: 10.1016/j.isci.2024.109713.
12. Omar M, Nadkarni GN, Klang E, Glicksberg BS. Large language models in medicine: A review of current clinical trials across healthcare applications. PLOS Digit Health. 2024;3(11):e0000662. doi: 10.1371/journal.pdig.0000662.
13. OpenAI, Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, et al. GPT-4 Technical Report. arXiv. 2023. doi: 10.48550/arXiv.2303.08774.
14. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. JAMA Intern Med. 2023;183(6):589-596. doi: 10.1001/jamainternmed.2023.1838.
15. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29:1930-1940. doi: 10.1038/s41591-023-02448-8.
16. Ho CN, Tian T, Ayers AT, Aaron RE, Phillips V, Wolf RM, Mathioudakis N, Dai T, Klonoff DC. Qualitative metrics from the biomedical literature for evaluating large language models in clinical decision-making: a narrative review. BMC Med Inform Decis Mak. 2024;24(1):357. doi: 10.1186/s12911-024-02757-z.
17. Lee J, Park S, Shin J, Cho B. Analyzing evaluation methods for large language models in the medical field: a scoping review. BMC Med Inform Decis Mak. 2024;24:366. doi: 10.1186/s12911-024-02709-7.
18. Guo Y, Ovadje A, Al-Garadi MA, Sarker A. Evaluating large language models for health-related text classification tasks with public social media data. J Am Med Inform Assn. 2024;31:2181-2189. doi: 10.1093/jamia/ocae210.
19. Tang YD, Dong ED, Gao W. LLMs in medicine: The need for advanced evaluation systems for disruptive technologies. Innovation (Camb). 2024;5(3):100622. doi: 10.1016/j.xinn.2024.100622.

Research Article

Comprehensive analysis and correlation of Serum 25-Hydroxyvitamin D3 Levels with glycemia and cardiometabolic risk factors: A tertiary care hospital-based study

Amit D. Sonagra[1], Deepak N Parchwani[2*], Anita Motiani[1], Sagar J. Dholariya[1], Ragini Singh[3], Krupal J Joshi[4]

[1]*Associate Professor (Biochemistry); All India Institute of Medical Sciences, Rajkot (Guj) India*
[2]*Professor (Biochemistry); All India Institute of Medical Sciences, Rajkot (Guj) India*
[3]*Additional Professor (Biochemistry); All India Institute of Medical Sciences, Rajkot (Guj) India*
[4]*Additional Professor (Community and Family Medicine); All India Institute of Medical Sciences, Rajkot (Guj) India*

Article Info

[*]***Corresponding Author:***
Deepak Parchwani
Professor (Biochemistry); All India Institute of Medical Sciences, Rajkot (Guj) India
E-mail: drdeepakparchwani@yahoo.com, drdparchwani@gmail.com

Keywords

HbA1c, dyslipidemia, lipid profile, vitamin D deficiency

Abstract

Background: Role of Vitamin D is not limited to skeletal health but it acts as a potential modulator of metabolic homeostasis. Studies have shown its influence on insulin secretion and cardiometabolic risk factors. This study was conducted to evaluate correlation of serum 25-hydroxy vitamin D3 (25OH VitD3) levels with glycemic status and lipid profile parameters in a tertiary care hospital setting.

Methods: A cross-sectional observational study was conducted taking 173 adults (>18 years) attending outpatient department of a tertiary care hospital between April and December 2025. Participants were categorized based on HbA1c levels according to American Diabetes Association criteria into non-diabetic, pre-diabetic, and diabetic. 25OH VitD3, total cholesterol (TC), triglycerides (TG), high-density lipoprotein (HDL), and low-density lipoprotein (LDL) were estimated from fasting serum sample. Statistical analyses were performed and statistical significance was set at $p<0.05$.

Results: Vitamin D deficiency and insufficiency were observed in more than 65% of participants (mean 25OH VitD3: 27.8±16.16 ng/mL). Patients with diabetes mellitus showed significantly higher TG, VLDL, TG/HDL and TC/HDL ratios and lower HDL levels ($p<0.05$). 25OH VitD3 showed a weak statistically significant negative correlation with HbA1c ($\rho = –0.186$, $p=0.014$) but no significant correlation with lipid parameters. Regression models revealed moderate explanatory power (R^2: 23.9–28.4%), with classical metabolic markers contributing more than vitamin D.

Conclusion: A high prevalence of vitamin D deficiency was detected. Vitamin D demonstrated a weak correlation with glycemia and no correlation with lipid parameters. Vitamin D may function as a metabolic modulator rather than a primary determinant of dysglycemia and dyslipidemia.

Introduction

Exploration of the functional role of vitamin D beyond its classical involvement in calcium homeostasis and skeletal health has generated considerable interest in recent years. Traditionally recognized for its importance in bone mineralization and prevention of rickets and osteomalacia, vitamin D is now increasingly acknowledged as a pleiotropic hormone with wide-ranging effects on multiple organ systems [1]. The discovery of vitamin D receptors (VDRs) in diverse tissues - including the pancreas, liver, adipose tissue, vascular smooth muscle, and immune cells - has significantly expanded the understanding of its biological functions. The active form of vitamin D, calcitriol (1,25-dihydroxyvitamin D), exerts genomic as well as non-genomic effects through its interaction with VDRs, thereby influencing cellular proliferation, differentiation, immune modulation, and metabolic regulation [2].

Particularly noteworthy is the presence of VDRs in pancreatic beta cells, which play a central role in insulin synthesis and secretion. Experimental and clinical studies suggest that vitamin D may modulate insulin secretion directly by binding to VDRs in beta cells and indirectly by regulating intracellular calcium flux, which is essential for insulin exocytosis. In addition, vitamin D has been implicated in improving insulin sensitivity in peripheral tissues such as skeletal muscle and adipose tissue. It may also exert anti-inflammatory effects that help mitigate insulin resistance, a key pathogenic mechanism underlying type 2 diabetes mellitus [3,4]. Thus, vitamin D appears to contribute to metabolic homeostasis not only through its endocrine functions but also via its immunomodulatory and anti-inflammatory properties.

Diabetes mellitus is characterized by chronic hyperglycemia resulting from impaired insulin secretion, insulin resistance, or both. Persistent disturbances in glucose metabolism are frequently accompanied by dyslipidemia, which significantly enhances cardiovascular risk. In patients with diabetes, increased hepatic lipogenesis and altered lipid metabolism lead to elevated triglyceride levels and reduced high-density lipoprotein (HDL) cholesterol concentrations. This dyslipidemic pattern promotes the formation of small dense low-density lipoprotein (LDL) particles, which are more susceptible to oxidative modification and more atherogenic than larger LDL particles [5,6]. The coexistence of hyperglycemia and atherogenic dyslipidemia accelerates endothelial dysfunction, oxidative stress, and chronic low-grade inflammation, all of which contribute to the progression of atherosclerosis and heightened cardiometabolic risk [7].

Emerging evidence suggests that vitamin D deficiency may be associated with adverse glycemic status and unfavorable lipid profiles. Low serum 25-hydroxyvitamin D3 [25(OH) D3] levels have been linked to poor glycemic control, higher HbA1c levels, insulin resistance, and an increased prevalence of metabolic syndrome. Moreover, vitamin D deficiency has been correlated with elevated triglycerides and reduced HDL cholesterol, further amplifying cardiovascular risk [8]. However, the findings across different populations remain inconsistent, possibly due to variations in ethnicity, lifestyle, dietary habits, sunlight exposure, and genetic factors affecting vitamin D metabolism.

In the Indian context, vitamin D deficiency is highly prevalent despite abundant sunlight, owing to factors such as skin pigmentation, urban lifestyle, limited outdoor activity, and dietary insufficiency. Studies have reported significantly high prevalence of hypovitaminosis D across different regions of India [9–11]. Nevertheless, studies examining the relationship between serum 25-hydroxyvitamin D3 levels, glycemic indices, and lipid parameters in the Indian population are relatively limited. Given the growing burden of diabetes and cardiovascular disease in India, understanding these associations is of considerable clinical relevance.

Therefore, the present study was undertaken to analyze the correlation of serum 25-hydroxyvitamin D3 levels with glycated hemoglobin (HbA1c) and lipid profile parameters. By evaluating these relationships, the study aims to provide further insight into the potential role of vitamin D in modulating glycemic control and cardiometabolic risk factors among individuals with altered glucose metabolism.

Materials and Methods

A cross-sectional observational study was conducted in a tertiary care teaching hospital after obtaining approval from the Institutional Ethics Committee. A total of 173 patients aged more than 18 years who attended the outpatient department of the hospital attached to a medical sciences institute were enrolled between April 2025 and December 2025 after obtaining written informed consent. The study was carried out in accordance with the ethical principles outlined in the Declaration of Helsinki and its subsequent revisions concerning biomedical research involving human subjects.

Non-diabetic, pre-diabetic and patients with diabetes mellitus, were assigned with group 1, 2 & 3 respectively. Diagnosis of diabetic state was made based on the glycated hemoglobin (HbA1c) value according to American Diabetes Association (ADA) criteria [12]. HbA1c was measured in a high-performance liquid chromatography (HPLC) based fully automatic analyser (Arkray - ADAMS HA 8180T, Japan). Any patients who were on therapy of lipid lowering agents and/or vitamin D3 supplementation, patients who refused to give consent were excluded in the study.

Recruited participants were instructed for 12 hours overnight fast. In the morning, 04 mL of venous blood was collected in plain vacutainer tube with gel separator under aseptic precaution. Blood was allowed to clot for 30 minutes and serum was separated by centrifugation at 1200g for 15 minutes. Fasting serum samples were analysed for 25-Hydroxyvitamin D3 (25OH Vit D3), total cholesterol (TC), triglycerides (TG), low density lipoprotein (LDL) and high-density lipoprotein (HDL) in fully automated analyser (Vitros 5600 integrated

system, QuidelOrtho, USA) with intra- and inter-assay coefficients of variation (CV) of 5.1 & 6.8, 1.1 & 1.5, 1.0 & 1.8, 1.5 & 2.9, 1.0 & 2.1 for the parameters respectively. Patients' demographic and clinical details were noted. Data was compiled using Microsoft Excel 2019 (Microsoft Corp., Redmond, WA) and statistical analysis was performed using Jamovi for Windows, version 2.6.26. Normality was assessed using the Shapiro–Wilk test. Normally distributed variables were analysed using Student's t-test or ANOVA, while non-parametric variables were analysed using Mann–Whitney U test or Kruskal–Wallis test. Spearman's rank correlation was used for association analysis. Multivariate linear regression models were constructed to evaluate independent predictors. Statistical significance was set at p value <0.05.

Results

Table 1 shows descriptive statistics and normality assessment of the study participants' data. Values of age, TC, LDL and non-HDL cholesterol followed normal distribution while the other parameters such as HbA1c, HDL, TG, VLDL, TG/HDL ratio, TC/HDL ratio, LDL/HDL ratio and 25OH Vit D3 values did not follow the normal distribution pattern. Non-parametric statistical tests were used to analyse the data which did not follow the normal distribution [13].

Table 1: Descriptive statistics and assessment of normality of the data.

Parameters	N	Mean	Standard Deviation	Shapiro-Wilk Test	
				W	p
Age (Years)	173	47.94	16.25	0.984	0.046
HbA1c (%)	173	6.67	1.94	0.709	<.001
Serum Total Cholesterol (mg/dL)	173	180.62	45.87	0.988	0.133
Serum HDL (mg/dL)	173	45.91	15.38	0.842	<.001
Serum LDL (mg/dL)	173	116.27	42.08	0.984	0.052
Serum Triglycerides (mg/dL)	173	153.83	79.49	0.843	<.001
Serum VLDL (mg/dL)	173	30.91	16.58	0.816	<.001
Triglyceride to HDL Ratio	173	3.85	2.81	0.779	<.001
Total Cholesterol to HDL Ratio	173	4.21	1.4	0.94	<.001
LDL to HDL Ratio	173	2.74	1.2	0.957	<.001
Non-HDL Cholesterol (mg/dL)	173	134.71	43.67	0.985	0.066
Serum 25-OH Vitamin D3 (ng/mL)	173	27.8	16.16	0.921	<.001

Table 2 shows comparison of parameters between female and male study participants. We found statistically significant lower concentration of HDL in male compared with female (p<0.05) while higher levels of TG/HDL ratio, TC/HDL ratio and LDL/HDL ratio in male compared with female (p<0.05). Other parameters were not statistically different among male and female study participants (p>0.05).

Table 2: Comparison of study participants as per gender.

Parameters	Female (n=103)		Male (n=70)		p value
	Mean / Median	SD / IQR	Mean / Median	SD / IQR	
Age (Years)*	46.99	15.15	49.34	17.77	0.352
HbA1c (%)†	5.9	1.05	6.1	1.8	0.11
Serum Total Cholesterol (mg/dL) *	181.82	45.75	178.86	46.32	0.678
Serum HDL (mg/dL)†	48	17.5	40	11.75	<0.001
Serum LDL (mg/dL) *	115.92	41.83	116.78	42.75	0.896
Serum Triglycerides (mg/dL)†	129	65	143.5	83.96	0.185

Parameters	Female (n=103)		Male (n=70)		p value
	Mean / Median	SD / IQR	Mean / Median	SD / IQR	
Serum VLDL (mg/dL)†	26	13.1	28.5	19.65	0.185
Triglyceride to HDL Ratio†	2.64	2.48	3.63	2.84	0.008
Total Cholesterol to HDL Ratio†	3.74	1.64	4.37	1.88	0.002
LDL to HDL Ratio†	2.38	1.55	2.85	1.52	0.009
Non-HDL Cholesterol (mg/dL) *	134	60.05	133	67.25	0.538
Serum 25-OH Vitamin D3 (ng/mL)†	23.7	19.95	23.4	21.92	0.924

**For parameters following normal distribution: Mean, SD and p value of student's t-test;*

† For parameters not following normal distribution: Median, IQR and p value of Mann-Whitney U test.

Table 3 shows comparison of parameters between non-diabetic, pre-diabetic and diabetic study participants classified as per ADA criteria. We found statistically significant difference in age, HbA1c, HDL, TG, VLDL, TG/HDL ratio, TC/HDL ratio and non-HDL cholesterol values among study groups ($p<0.05$). Other parameters did not show statistically significant difference.

Table 3: Comparison of study participants based on the diabetic status.

Parameters	Non-Diabetic (n=53)		Pre-Diabetic (n=63)		Diabetic (n=57)		p value
	Mean / Median	SD / IQR	Mean / Median	SD / IQR	Mean / Median	SD / IQR	
Age (Years) *	36.45	16.19	47.81	14.20	58.77	10.01	<0.001
HbA1c (%)†	5.5	0.2	5.9	0.3	7.8	3.3	<0.001
Serum Total Cholesterol (mg/dL) *	171.45	39.12	189.98	42.32	178.79	53.64	0.053
Serum HDL (mg/dL)†	47	13	46	14.5	40	18	0.006
Serum LDL (mg/dL) *	108.48	37.33	123.54	38.68	115.48	48.69	0.109
Serum Triglycerides (mg/dL)†	111	92	135	67	146	78	0.004
Serum VLDL (mg/dL)†	22	18.4	27	13.4	29.2	16	0.004
Triglyceride to HDL Ratio†	2.28	2.61	2.88	1.96	4.03	2.62	<0.001
Total Cholesterol to HDL Ratio †	3.61	1.19	3.92	1.62	4.71	1.92	0.004
LDL to HDL Ratio†	2.38	1.14	2.56	1.39	3.01	1.88	0.065
Non-HDL Cholesterol (mg/dL) *	121.74	39.25	143.08	40.25	137.53	48.86	0.016
Serum 25-OH Vitamin D3 (ng/mL)†	23.7	21	24	19.1	22.8	18.4	0.63

**For parameters following normal distribution: Mean, SD and p value of ANOVA test;*

† For parameters not following normal distribution: Median, IQR and p value of Kruskal-Wallis test.

Table 4 shows comparison of parameters between participants having deficiency, insufficiency and sufficiency as per serum 25OH VitD3 levels. There was no statistically significant difference in all the parameters ($p>0.05$) except 25OH VitD3 levels ($p<0.05$) among study groups.

Table 4: Comparison of study participants based on the Serum 25-OH Vitamin D3 levels.

Parameters	Deficient (<20 ng/mL) (n=62)		Insufficient (20-30 ng/mL) (n=52)		Sufficient (30-100 ng/mL) (n=59)		p value
	Mean / Median	SD / IQR	Mean / Median	SD / IQR	Mean / Median	SD / IQR	
Age (Years)*	45.66	16.45	48.67	17.4	49.69	14.93	0.357
HbA1c (%)†	5.8	2.56	5.95	1.15	5.9	1.15	0.855
Serum Total Cholesterol (mg/dL)*	172.42	46.39	189.54	45.89	181.37	44.51	0.147
Serum HDL (mg/dL)†	41	19.75	46.5	14.25	45	10	0.202
Serum LDL (mg/dL)*	109.51	40.07	121.26	44.19	118.98	42.03	0.271
Serum Triglycerides (mg/dL)†	133.5	85.75	142	78.5	129	77.5	0.339
Serum VLDL (mg/dL)†	26.5	17.15	28.2	15.7	26	15.5	0.338
Triglyceride to HDL Ratio†	2.79	3.22	3.53	2.42	2.93	2.04	0.694
Total Cholesterol to HDL Ratio †	3.99	1.65	4.17	1.81	3.92	1.64	0.643
LDL to HDL Ratio†	2.56	1.44	2.86	1.26	2.45	1.59	0.912
Non-HDL Cholesterol (mg/dL)*	129.4	44.92	141.42	45.69	134.37	40.34	0.375
Serum 25-OH Vitamin D3 (ng/mL)†	13.75	5.47	23.6	4.7	43.2	13.85	<0.001

**For parameters following normal distribution: Mean, SD and p value of ANOVA test;*

† For parameters not following normal distribution: Median, IQR and p value of Kruskal-Wallis test.

Table 5 shows Spearman's correlation matrix heatmap among parameters. We found significant positive correlation between age vs HbA1c, age vs TG, age vs VLDL, age vs TG/HDL ratio, HbA1c vs TG, HbA1c vs VLDL, HbA1c vs TG/HDL ratio, HbA1c vs TC/HDL ratio, HbA1c vs LDL/HDL ratio, TC vs HDL, TC vs LDL, TC vs TG, TC vs VLDL, TC vs TC/HDL ratio, TC vs LDL/HDL ratio, TC vs non-HDL, LDL vs TG, LDL vs VLDL, LDL vs TC/HDL ratio, LDL vs LDL/HDL ratio, LDL vs non-HDL, TG vs VLDL, TG vs TG/HDL ratio, TG vs TC/HDL ratio, TG vs LDL/HDL ratio, TG vs non-HDL, VLDL vs TG/HDL ratio, VLDL vs TC/HDL ratio, VLDL vs LDL/HDL ratio, VLDL vs non-HDL, TG/HDL ratio vs TC/HDL ratio, TG/HDL ratio vs LDL/HDL ratio, TG/HDL ratio vs non-HDL, TC/HDL ratio vs LDL/HDL ratio, TC/HDL ratio vs non-HDL and LDL/HDL ratio vs non-HDL ($p<0.05$). We found significant negative correlation between HbA1c vs HDL, HbA1c vs 25OH VitD3 (also shown in Figure 1), HDL vs TG, HDL vs VLDL, HDL vs TG/HDL ratio, HDL vs TC/HDL ratio and HDL vs LDL/HDL ratio ($p<0.05$). Remaining correlations were not statistically significant.

Table 5: Spearman's correlation matrix heatmap among parameters (n=173).

Parameters		Age	HbA1c	TC	HDL	LDL	TG	VLDL	TG/HDL	TC/HDL	LDL/HDL	Non-HDL	25OHD
Age	rho (ρ)	—											
	p-value	—											
HbA1c	rho (ρ)	0.565*	—										
	p-value	<.001	—										
TC	rho (ρ)	0.041	0.028	—									
	p-value	0.593	0.714	—									
HDL	rho (ρ)	-0.055	-0.266*	0.297*	—								
	p-value	0.47	<.001	<.001	—								
LDL	rho (ρ)	0.013	0.025	0.890*	0.094	—							
	p-value	0.862	0.741	<.001	0.219	—							
TG	rho (ρ)	0.160*	0.256*	0.310*	-0.305*	0.224*	—						
	p-value	0.036	<.001	<.001	<.001	0.003	—						
VLDL	rho (ρ)	0.157*	0.257*	0.309*	-0.305*	0.224*	1.000*	—					
	p-value	0.039	<.001	<.001	<.001	0.003	<.001	—					
TG/HDL Ratio	rho (ρ)	0.157*	0.329*	0.087	-0.676*	0.122	0.890*	0.890*	—				
	p-value	0.04	<.001	0.256	<.001	0.11	<.001	<.001	—				
TC/HDL Ratio	rho (ρ)	0.102	0.281*	0.499*	-0.617*	0.626*	0.535*	0.535*	0.690*	—			
	p-value	0.183	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—			
LDL/HDL Ratio	rho (ρ)	0.061	0.201*	0.537*	-0.525*	0.750*	0.383*	0.383*	0.534*	0.947*	—		
	p-value	0.429	0.008	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—		
Non-HDL	rho (ρ)	0.072	0.112	0.940*	0.036	0.935*	0.416*	0.415*	0.292*	0.717*	0.740*	—	
	p-value	0.346	0.144	<.001	0.637	<.001	<.001	<.001	<.001	<.001	<.001	—	
25OH VitD3	rho (ρ)	0.062	-0.186*	0.095	0.146	0.096	0.021	0.022	-0.056	-0.086	-0.05	0.069	—
	p-value	0.415	0.014	0.214	0.055	0.208	0.779	0.777	0.466	0.263	0.515	0.367	—

**Statistically significant correlation.*

Figure 1: Scatter plot of 25OH VitD3 vs HbA1c among study participants (n=173).

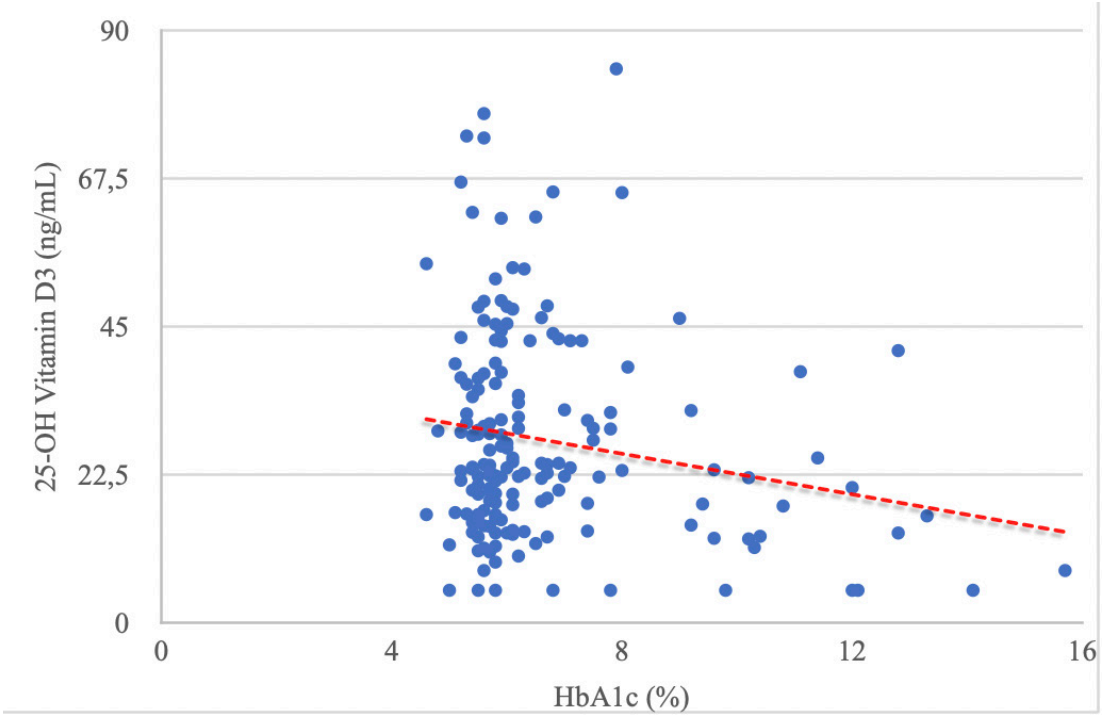


Table 6 to 9 show the statistics of linear regression model that was developed taking different predictors and outcome variables. Each model had a moderate positive relationship between the predictor variables and outcome variables as described by the correlation coefficient (R) and R-squared values. The mean Cook's Distance value for each model was quite low, suggesting that the influence of individual data points on the estimated regression coefficients was minimal. This indicates that there were no influential outliers that could significantly distort the results [14]. All the models were statistically significant ($p<0.05$).

Table 6: Linear Regression model statistics: Model 1.

Model	R	R^2	Adjusted R^2	Overall Model Test				Mean Cook's Distance
				F	df1	df2	p	
1	0.489	0.239	0.211	8.68	6	166	<0.001	0.00871

Note: *1. Model estimated using sample size of N=173*
2. Dependent / Outcome Variable: HbA1c
3. Independent / Predictor Variables: Age, Gender, TC, TG, HDL, 25OH VitD3

Table 7: Linear Regression model statistics: Model 2.

Model	R	R^2	Adjusted R^2	Overall Model Test				Mean Cook's Distance
				F	df1	df2	p	
2	0.511	0.262	0.235	9.8	6	166	<0.001	0.0112

Note: *1. Model estimated using sample size of N=173*
2. Dependent / Outcome Variable: TC
3. Independent / Predictor Variables: Age, Gender, HbA1c, TG, HDL, 25OH VitD3

Table 8: Linear Regression model statistics: Model 3.

Model	R	R^2	Adjusted R^2	Overall Model Test				Mean Cook's Distance
				F	df1	df2	p	
3	0.492	0.242	0.214	8.83	6	166	<0.001	0.00849

Note: *1. Model estimated using sample size of N=173*
2. Dependent / Outcome Variable: TG
3. Independent / Predictor Variables: Age, Gender, HbA1c, TC, HDL, 25OH VitD3

Table 9: Linear Regression model statistics: Model 4.

Model	R	R^2	Adjusted R^2	Overall Model Test				Mean Cook's Distance
				F	df1	df2	p	
4	0.533	0.284	0.258	11	6	166	<0.001	0.00658

Note: *1. Model estimated using sample size of N=173*
2. Dependent / Outcome Variable: HDL
3. Independent / Predictor Variables: Age, Gender, HbA1c, TC, TG, 25OH VitD3

Discussion

The present tertiary care hospital-based study provided a comprehensive evaluation of 25OH VitD3 levels in relation to glycemic status and cardiometabolic risk factors. Our finding will contribute to resolve ongoing debate regarding whether vitamin D deficiency is only a marker of poor metabolic health or it is an independent determinant of dysglycemia and dyslipidemia.

One of the important observations from the study stated that more than 35% study participant were vitamin D3 deficient and more than 30% were having insufficient levels of vitamin D3. Mean serum level of 25OH VitD3 were 27.8 ng/mL (n=173) which again falls within vitamin D3 insufficiency range that depicts the magnitude of the subnormal vitamin D status. Many other studies have also reported similar high prevalence of vitamin D deficiency even in apparently healthy populations [15–19]. Despite India being a sun-light rich country, such paradoxically low vitamin D levels may be due to adaptation to urban lifestyle, indoor occupational pattern, reduced outdoor physical activity, darker skin pigmentation, dietary insufficiency, and pollution-related UVB attenuation can be few of the major factors causing hypovitaminosis D [16,20,21]. Upon normality check with Shapiro-Wilk Test, we found many metabolic parameters did not follow a Gaussian distribution. This suggests heterogenous distribution of these cardiometabolic parameters in hospital-based population. To maintain appropriateness, strength and validity of the statistical inferences, non-parametric tests were utilized for the variables not following normal distribution [13].

Gender based differences in lipid metabolism exist due to difference in hormonal milieu between them. Oestrogen enhances hepatic LDL receptor expression, promotes reverse cholesterol transport and increases apolipoprotein A1 synthesis that increases HDL concentration. On the other side, greater visceral obesity along with androgens are associated increased flux of free fatty acids into portal circulation enhancing hepatic VLDL synthesis. Androgens also increase hepatic lipase activity which along high TG increases HDL catabolism promoting atherogenic lipid profile pattern [22]. In our study, male participants exhibited significantly lower HDL and higher TG/HDL, TC/HDL and LDL/HDL ratios compared to females. These findings are consistent with previously done studies [22–24]. However, 25OH VitD3 did not differ between both the genders, which suggests that gender-based difference in lipid indices may not be attributed to vitamin D status of the individuals.

A comparison of cardiometabolic indices was made in the study participants divided in three groups based on ADA defined glycemic categories. We found significant difference in age, HDL, TG, VLDL, TG/HDL ratio, TC/HDL ratio and non-HDL cholesterol. Diabetic individuals were relatively older in age and they exhibited atherogenic dyslipidemia. They had elevated TG, VLDL, TG/HDL and TC/HDL ratios and lower HDL levels compared with non-diabetic participants. These findings show characteristic pattern of diabetic dyslipidemia caused by insulin resistance [25,26]. Insulin inhibits hormone sensitive lipase and reduces free fatty acid flux to the liver and thereby reducing hepatic VLDL production. In the state of insulin resistance (IR), there is increased lipolysis in adipose tissue resulting increased free fatty delivery to hepatocyte and causing increased TG synthesis and VLDL secretion. IR also reduce lipoprotein lipase activity and impair peripheral TG clearance. Elevated TG increases cholesterol ester transfer protein (CETP) mediated exchange of TG and cholesterol ester between HDL and TG rich VLDL causing HDL to become TG rich and making them more susceptible to hepatic lipase mediated catabolism and leading to reduced serum HDL level [27]. Many studies have shown elevated TG/HDL ratio as a surrogate marker of insulin resistance and increased levels of small dense LDL particles which have shown to participate in

atherogenic plaque formation in blood vessels [27–29]. However, serum 25OH Vit D3 did not significantly differ across these groups divided as per glycemic status, which suggest that despite high prevalence of vitamin D deficiency, it may not be a primary determinant for glycemic dysregulation and dyslipidemia in the study participants. We also stratified the participants as per their vitamin D status as deficient, insufficient, and sufficient groups. No statistically significant differences were observed in glycemic indices, lipid parameters and atherogenic ratios. This finding further reinforces the fact that vitamin D deficiency alone is not the driving force for cardiometabolic derangements in clinical settings [30].

Analysis of correlation matrix revealed many significant relations between various cardiometabolic parameters. A strong positive correlation between age and HbA1c suggest age as a strong determinant for dysglycemic state. Such changes can be attributed to occurrence of gradual β-cell dysfunction, reduced incretin effect, mitochondrial functional decline, sarcopenia, reduced physical activity, increased occurrence of obesity and chronic low-grade inflammation as the age advances [31]. Correlation of HbA1c with lipid profile and ratios further reflects bidirectional relation between dyslipidemia and insulin resistance. Hepatic de-novo lipogenesis is activated by hyperglycemia through activation of carbohydrate-responsive element-binding protein (ChREBP) and sterol regulatory element-binding protein-1c (SREBP-1c) which increases TG synthesis. On the other side, lipid accumulation in skeletal muscle and liver impairs insulin signalling via diacylglycerol (DAG) mediated activation of protein kinase C (PKC) pathways. Both these mechanisms establish vicious cycle and enhance effects of each other [28,32,33].

We observed weak but a significant negative correlation ($\rho = -0.186$) of 25OH VitD3 with HbA1c. Vitamin D did not show any significant correlation with other lipid indices. Vitamin D have found to affect metabolism in several way. The pancreatic β-cells expresses vitamin D receptors (VDR). Vitamin D is found to enhance insulin gene transcription through vitamin D response elements. It is also associated with modulation of intracellular calcium flux that drives the insulin exocytosis. Thus, deficiency of vitamin D may impair insulin secretion and affect glucose homeostasis [2]. On the other side, vitamin D reduces pro-inflammatory cytokines (TNF-α, IL-6) that interfere with insulin signalling by inducing serine phosphorylation instead of tyrosine phosphorylation of signalling pathway intermediates. Chronic vitamin D deficiency may promote low-grade inflammation and insulin resistance [31,34]. Vitamin D also influences adipocyte differentiation and lipid storage that further affect adipose tissue homeostasis [35]. We could not find any direct significant correlation of vitamin D with lipid profile indices. However, strong correlations of HbA1c with lipid profile and negative correlation of Vitamin D with HbA1c can indirectly suggest effect of vitamin D on person's lipid profile by modulating the metabolism as mentioned above. Findings from our study aligns with the findings reported by Liu et al [8], Ge L et al [36], Alqahtani RM et al [37] and Abdo B et al [38]. Studies have also suggested that vitamin D supplementation in vitamin D deficient persons do not affect the dyslipidemia in them [36,39–41]. Many researchers have found a significant association of dyslipidemia with vitamin D deficiency [42–46]. Though similar pattern of lipid profile changes was observed, the changes are not statistically significant in our study. Mechanism for changes in lipid profile in vitamin D deficiency can be increased cholesterol synthesis via increased 3-hydroxy-3-methylglutaryl-coenzyme A (HMG CoA) reductase expression through insulin-induced gene-2 (Insig-2) expression down-regulation which is followed by inhibition of sterol regulatory element-binding protein 2 (SREBP-2) through reduced transcriptional activity of Vitamin D receptor (VDR) in vitamin D deficiency [47]. Vitamin D also modulate microsomal triglyceride transfer protein (MTP), and thus reduces the synthesis and secretion of triglycerides [48]. Vitamin D also affects insulin signalling and inflammatory state, thereby affects serum TG and HDL levels [49]. Thus, correlations of HbA1c with vitamin D and lipid profile parameters in our study suggests that vitamin D may act as a secondary modulator rather than a primary driver of hyperglycaemia by modestly influencing the β-cell responsiveness and insulin receptor signalling, thereby influencing lipid metabolism [38].

Four regression models were developed to explore independent determinants. Each model had a moderate R value and low mean Cook's distance that confirmed model's stability and absence of influential outliers. Collectively all regression models highlight that HbA1c and lipid indices accounted much for explaining variances in them along with age and gender while 25OH VitD3 contributed minimally. Analysis further stats that vitamin D though has an important role in metabolism, it is not the only and primary determinant for development of dysglycemia and dyslipidemia [30].

Cross-sectional observational studies have several limitations that should be considered when interpreting the results. We could not establish causal relationship between glycemic and cardiometabolic parameters with vitamin D deficiency. Longitudinal studies are needed to identify these associations. Certain unmeasured factors such as genetic predispositions, dietary habits, physical activity, body mass index, central obesity indices, addictions, seasonal variation in vitamin D synthesis, etc, can substantially influence the correlations. Being a single centre study with limited number of participants, the study results cannot be implemented to general population. This study provides a future direction for planning of a multicentric longitudinal study with participant of different ethnicity, geographical regions and socioeconomic groups that would strengthen the findings and would make them more reliable and applicable in different populations. A randomized controlled trials conducted with adequate dosing and specified

duration may help conclude whether correction of vitamin D deficiency confers metabolic benefit in specific subgroups.

Conclusion

Vitamin D deficiency and insufficiency is highly prevalent in the population. Screening and supplementation for them can improve overall skeletal health. Serum 25OH VitD3 had weak negative correlation with HbA1c and it was not a strong independent predictor in multivariate regression analysis. Classical metabolic markers remain stronger determinants of cardiometabolic risk. Thus, vitamin D may serve as a contributory metabolic modulator rather than a primary etiological factor in dysglycaemia and dyslipidemia. Vitamin D correction cannot be viewed as a standalone strategy for glycaemic and lipemic control.

Competing interests

The authors declare that they have no competing interests to disclose.

Declaration of conflict of interest

The authors affirm that this study is free from any conflicts of interest.

Ethical Approval

This investigation was conducted in accordance with the ethical guidelines of the Declaration of Helsinki on biomedical research on humans and was approved by the Institutional Ethics Committee of AIIMS Rajkot. Letter Number: AIIMS/RAJKOT/IEC/6th/ER/07.

Funding Statement

The authors received no financial support for the research, authorship, and/or publication of this article.

Data Availability Statement

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Credit Authors:

[1]AS: Conceptualization, Methodology, Review, Editing and validation, Data curation, supervision and editing.
[2]DP: Conceptualization, Methodology, Review, Editing and validation.
[1]AM: Data curation, supervision and editing.
[1]SD: Original Draft, Data curation, editing.
[2]RS: Original draft, software, data curation, and review.
[3]KJ: Original draft, software, data curation, and review.

The corresponding author affirms that all listed authors have reviewed and approved their respective contributions and agree with the final version of the manuscript for submission.

Acknowledgment

This paper is based on Departmental study No. AIIMS/RAJKOT/IEC/6th/ER/07 granted by All India Institute of Medical Sciences, Rajkot.

Supplier name for reagents/apparatus

1. Reagents/consumables/column for HPLC based ArkRay ADAMS- HA 8180T for HbA1c: Authorized distributor - SPSS Medline; New Delhi (India)
2. Fully automatic integrated chemistry and Immunoassay analyzer: Vitros 5600; QuidelOrtho USA
3. Kits/reagents/consumables for Vitro 5600 integrated analyzer: Authorized distributor – Olympus, Rajkot (India)
4. ArkRay ADAMS- HA 8180T HPLC analyzer for HbA1c: Arkray Japan
5. Control: Biorad; Genetix Ahmedabad (India)

References:

1. Voltan G, Cannito M, Ferrarese M, Ceccato F, Camozzi V. Vitamin D: An Overview of Gene Regulation, Ranging from Metabolism to Genomic Effects. Genes (Basel). 2023;14(9). doi:10.3390/GENES14091691
2. Argano C, Mirarchi L, Amodeo S, Orlando V, Torres A, Corrao S. The Role of Vitamin D and Its Molecular Bases in Insulin Resistance, Diabetes, Metabolic Syndrome, and Cardiovascular Disease: State of the Art. Int J Mol Sci. 2023;24(20). doi:10.3390/IJMS242015485
3. Dong B, Zhou Y, Wang W, et al. Vitamin D Receptor Activation in Liver Macrophages Ameliorates Hepatic Inflammation, Steatosis, and Insulin Resistance in Mice. Hepatology. 2020;71(5):1559-1574. doi:10.1002/HEP.30937
4. Zhang Y, Ni P, Miao Y, et al. Vitamin D3 improves glucose metabolism and attenuates inflammation in prediabetic human and mice. J Nutr Biochem. 2024;130. doi:10.1016/J.JNUTBIO.2024.109659
5. Wong ND, Sattar N. Cardiovascular risk in diabetes mellitus: epidemiology, assessment and prevention. Nat Rev Cardiol. 2023;20(10):685-695. doi:10.1038/S41569-023-00877-Z
6. Pennells L, Kaptoge S, Østergaard HB, et al. SCORE2-Diabetes: 10-year cardiovascular risk estimation in type 2 diabetes in Europe. Eur Heart J. 2023;44(28):2544-2556. doi:10.1093/EURHEARTJ/EHAD260
7. Hariharan R, Odjidja EN, Scott D, et al. The dietary inflammatory index, obesity, type 2 diabetes, and cardiovascular risk factors and diseases. Obes Rev. 2022;23(1). doi:10.1111/OBR.13349
8. Liu YJ, Duan JW, Lu DH, Zhang F, Liu HL. Association between vitamin D status and cardiometabolic risk factors in adults with type 2 diabetes in Shenzhen, China. Front Endocrinol (Lausanne). 2024;15. doi:10.3389/FENDO.2024.1346605
9. Bhaduri G, Samajdar SS, Chakraborty S, Saboo B, Joshi

SR. Prevalence and risk factors of hypovitaminosis D among patients with diabetes mellitus from India: A systematic review and meta-analysis. Diabetes Metab Syndr Clin Res Rev. 2025;19(10). doi:10.1016/j.dsx.2025.103326
10. Kamboj P, Dwivedi S, Toteja GS. Prevalence of hypovitaminosis D in India & way forward. Indian J Med Res. 2018;148(5):548-556. doi:10.4103/IJMR.IJMR_1807_18
11. Shivane VK, Sarathi V, Bandgar T, Menon P, Shah NS. High prevalence of hypovitaminosis D in young healthy adults from the western part of India. Postgrad Med J. 2011;87(1030):514-518. doi:10.1136/PGMJ.2010.113092
12. Committee ADAPP. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2025. Diabetes Care. 2024;48(Suppl 1):S27. doi:10.2337/dc25-S002
13. Vrbin CM. Parametric or nonparametric statistical tests: Considerations when choosing the most appropriate option for your data. Cytopathology. 2022;33(6):663-667. doi:10.1111/cyt.13174
14. Zhu H, Ibrahim JG, Cho H. PERTURBATION AND SCALED COOK'S DISTANCE. Ann Stat. 2012;40(2):785-811. doi:10.1214/12-AOS978
15. Kumar GT, Chugh R, Eggersdorfer M. Poor Vitamin D Status in Healthy Populations in India: A Review of Current Evidence. Int J Vitam Nutr Res. 2015;85(3-4):185-201. doi:10.1024/0300-9831/a000228
16. Ghosh A, Monisha S, Sunny AS, Diwakar L, Issac TG. Prevalence and patterns of vitamin D deficiency and its role in cognitive functioning in a cohort from South India. Sci Rep. 2024;14(1). doi:10.1038/s41598-024-62010-5
17. Karibayeva I, Bilibayeva G, Yerzhanova A, et al. Prevalence of Vitamin D Deficiency Among Adults in Kazakhstan: A Systematic Review and Meta-Analysis. Medicina (Kaunas). 2024;60(12). doi:10.3390/medicina60122043
18. Mendes MM, Gomes APO, Araújo MM, Coelho ASG, Carvalho KMB, Botelho PB. Prevalence of vitamin D deficiency in South America: a systematic review and meta-analysis. Nutr Rev. 2023;81(10):1290-1309. doi:10.1093/nutrit/nuad010
19. Mogire RM, Mutua A, Kimita W, et al. Prevalence of vitamin D deficiency in Africa: a systematic review and meta-analysis. Lancet Glob Heal. 2020;8(1):e134-e142. doi:10.1016/S2214-109X(19)30457-7
20. P A, S M, B N, SK G. Vitamin D deficiency in India. J Fam Med Prim care. 2018;7(2):41-44+55. doi:10.4103/jfmpc.jfmpc_78_18
21. Praveen PA, Singh A, Lakshmy R, et al. Prevalence and Correlates of Vitamin D Deficiency among Adult Population in Urban and Rural Areas of the National Capital Region of Delhi, India. WHO South-East Asia J public Heal. 2023;12(2):104-109. doi:10.4103/WHO-SEAJPH.WHO-SEAJPH_113_22
22. Palmisano BT, Zhu L, Eckel RH, Stafford JM. Sex differences in lipid and lipoprotein metabolism. Mol Metab. 2018;15(19):45-55. doi:10.1016/j.molmet.2018.05.008
23. Kolovou GD, Anagnostopoulou KK, Damaskos DS, et al. Gender differences in the lipid profile of dyslipidemic subjects. Eur J Intern Med. 2009;20(2):145-151. doi:10.1016/j.ejim.2008.06.011
24. Victor BM, Teal V, Ahedor L, Karalis DG. Gender differences in achieving optimal lipid goals in patients with coronary artery disease. Am J Cardiol. 2014;113(10):1611-1615. doi:10.1016/j.amjcard.2014.02.018
25. Ndzie Noah ML, Adzika GK, Mprah R, Adekunle AO, Adu-Amankwaah J, Sun H. Sex-Gender Disparities in Cardiovascular Diseases: The Effects of Estrogen on eNOS, Lipid Profile, and NFATs During Catecholamine Stress. Front Cardiovasc Med. 2021;8. doi:10.3389/fcvm.2021.639946
26. Ormazabal V, Nair S, Elfeky O, Aguayo C, Salomon C, Zuñiga FA. Association between insulin resistance and the development of cardiovascular disease. Cardiovasc Diabetol. 2018;17(1). doi:10.1186/s12933-018-0762-4
27. Alcover S, Ramos-Regalado L, Girón G, Muñoz-García N, Vilahur G. HDL-Cholesterol and Triglycerides Dynamics: Essential Players in Metabolic Syndrome. Antioxidants. 2025;14(4):434. doi:10.3390/antiox14040434
28. Duan M, Zhao X, Li S, et al. Metabolic score for insulin resistance (METS-IR) predicts all-cause and cardiovascular mortality in the general population: evidence from NHANES 2001-2018. Cardiovasc Diabetol. 2024;23(1). doi:10.1186/s12933-024-02334-8
29. Oliveri A, Rebernick RJ, Kuppa A, et al. Comprehensive genetic study of the insulin resistance marker TG:HDL-C in the UK Biobank. Nat Genet. 2024;56(2):212-221. doi:10.1038/s41588-023-01625-2
30. Che B, Zhong C, Zhang R, et al. Triglyceride-glucose index and triglyceride to high-density lipoprotein cholesterol ratio as potential cardiovascular disease risk factors: an analysis of UK biobank data. Cardiovasc Diabetol. 2023;22(1). doi:10.1186/s12933-023-01762-2
31. Manousaki D, Harroud A, Mitchell RE, et al. Vitamin D levels and risk of type 1 diabetes: A Mendelian randomization study. PLoS Med. 2021;18(2). doi:10.1371/JOURNAL.PMED.1003536
32. Uddin MF, Garapati A, Gyanendra KC, et al. Relationship Between Glycosylated Hemoglobin and Serum Lipid Profile in Patients With Type 2 Diabetes Mellitus: A Cross Sectional Analysis. Heal Sci Reports. 2025;8(7):e70965. doi:10.1002/hsr2.70965
33. Iizuka K, Takao K, Yabe D. ChREBP-Mediated Regulation of Lipid Metabolism: Involvement of the Gut Microbiota, Liver, and Adipose Tissue. Front Endocrinol (Lausanne). 2020;11:587189. doi:10.3389/fendo.2020.587189

34. Liu GT, Fenercioglu AK. The Anti-Inflammatory Roles of Vitamin D for Improving Human Health. Curr Issues Mol Biol 2024, Vol 46, Pages 13514-13525. 2024;46(12):13514-13525. doi:10.3390/cimb46120807
35. Lu S, Cao ZB. Interplay between Vitamin D and Adipose Tissue: Implications for Adipogenesis and Adipose Tissue Function. Nutr 2023, Vol 15, Page 4832. 2023;15(22):4832. doi:10.3390/nu15224832
36. Ge L, Gao S, Kia N, Wang Y, Hua L. Effect of vitamin D supplementation on body composition, lipid profile, and glycemic indices in patients with obesity-associated metabolic syndrome: a systematic review and meta-analysis. Diabetol Metab Syndr. 2025;17(1):287. doi:10.1186/S13098-025-01799-1
37. Alqahtani RM, Alsulami EF. The Association Between Glycated Hemoglobin (HbA1c) Level and Vitamin D Level in Diabetes Mellitus Patients: A Cross-Sectional Study. Cureus. 2023;15(10). doi:10.7759/CUREUS.47166
38. Abdo B, Abdullah M, AlShoaibi IA, et al. Relationship Between Glycated Hemoglobin (HbA1c) and Vitamin D Levels in Type 2 Diabetes Patients: A Retrospective Cross-Sectional Study. Cureus. 2024;16(6):e62468. doi:10.7759/CUREUS.62468
39. Lu Q, Liang Q, Xi Y. The effects of vitamin D supplementation on serum lipid profiles in people with type 2 diabetes: a systematic review and meta-analysis of randomized controlled trials. Front Nutr. 2024;11:1419747. doi:10.3389/FNUT.2024.1419747/FULL
40. Bahrami LS, Ranjbar G, Norouzy A, Arabi SM. Vitamin D supplementation effects on the clinical outcomes of patients with coronary artery disease: a systematic review and meta-analysis. Sci Reports 2020 101. 2020;10(1):12923-. doi:10.1038/s41598-020-69762-w
41. Rezaei S, Tabrizi R, Nowrouzi-Sohrabi P, et al. The Effects of Vitamin D Supplementation on Anthropometric and Biochemical Indices in Patients With Non-alcoholic Fatty Liver Disease: A Systematic Review and Meta-analysis. Front Pharmacol. 2021;12:732496. doi:10.3389/FPHAR.2021.732496/TEXT
42. Frențușcă C, Babeș K. Association between vitamin D deficiency and serum lipid levels in a group of Romanian patients. Med Pharm reports. 2024;97(3):263-269. doi:10.15386/MPR-2714
43. Cheng YL, Lee TI, Chien YM, Lee TW, Chen YJ. Vitamin D level regulates serum lipids discrepantly in adults with and without dyslipidemia. Endocr Connect. 2023;12(9). doi:10.1530/EC-23-0013
44. Sharba ZF, Shareef RH, Abd BA, Hameed EN. Association between Dyslipidemia and Vitamin D Deficiency: a Cross-Sectional Study. Folia Med (Plovdiv). 2021;63(6):965-969. doi:10.3897/FOLMED.63.E62417
45. Frentusca CF, Babes K, Galusca DI. VITAMIN D DEFICIENCY AS AN INDEPENDENT PREDICTOR OF CARDIOVASCULAR DISEASE. Acta Endocrinol (Bucharest, Rom 2005). 2023;19(3):319-325. doi:10.4183/AEB.2023.319
46. Li B, Wang J, Xu J, et al. Association between dyslipidemia and vitamin D deficiency: a cross-sectional study in Chinese healthy population. Front Endocrinol (Lausanne). 2025;16. doi:10.3389/FENDO.2025.1450924
47. Li S, He Y, Lin S, et al. Increase of circulating cholesterol in vitamin D deficiency is linked to reduced vitamin D receptor activity via the Insig-2/SREBP-2 pathway. Mol Nutr Food Res. 2016;60(4):798-809. doi:10.1002/MNFR.201500425
48. Hussain MM, Nijstad N, Franceschini L. Regulation of microsomal triglyceride transfer protein. Clin Lipidol. 2011;6(3):293-303. doi:10.2217/CLP.11.21
49. Malekmohammad K, Bezsonov EE, Rafieian-Kopaei M. Role of Lipid Accumulation and Inflammation in Atherosclerosis: Focus on Molecular and Cellular Mechanisms. Front Cardiovasc Med. 2021;8. doi:10.3389/FCVM.2021.707529

Research Article

Differential Effects of BMI Adjustment on Ferritin Levels in Anemic Patients

Dhivya Baskaran[1], Jagadish Ramasamy[1,2*]

[1]*Department of Biochemistry, Velammal Medical College Hospital and Research Institute, Madurai, Tamil Nadu, India*
[2]*Department of Biochemistry, All India Institute of Medical Sciences, Madurai, Tamil Nadu, India*

Article Info

[*]***Corresponding Author:***
Jagadish Ramasamy
Department of Biochemistry, Velammal Medical College Hospital and Research Institute, Madurai, Tamil Nadu, India
Current Affiliation: Department of Biochemistry, All India Institute of Medical Sciences, Madurai, Tamil Nadu, India
ORCID: 0000-0003-4725-3227
E-mail: iamjagankmr@gmail.com
Phone: +918015597917

Keywords

Ferritin, BMI, Obesity, regression, inflammation, Iron deficiency, anemia

Abstract

Background: The prevalence of obesity and anaemia is consistently increasing in India. Obesity is a chronic inflammatory state and is associated with dysregulation of iron homeostasis. Ferritin, a marker for cellular iron stores, is also an acute phase reactant that remains elevated during inflammation. Previous studies have used inflammation-adjusted serum ferritin to address this. This current study aims to calculate body mass index (BMI)-adjusted ferritin using a regression correction and evaluate its effects across gender and BMI category in anemic patients.

Methods: De-identified data were retrieved from hospital records and laboratory information systems. BMI-adjusted ferritin was estimated using a regression correction formula: ln(adjusted ferritin) = ln(unadjusted ferritin) − β(BMI − median BMI).

Results: A total of 458 subjects were included, of whom 62% were female and 95% were anemic. Adjusted ferritin levels were slightly lower than unadjusted levels in the normal BMI group (59.8 vs 60) and in males (102.3 vs 113.75). In contrast, adjusted ferritin levels were marginally higher in the overweight/obese group (29.9 vs 28.9) and in females (22.7 vs 22.1). The percentage change in ferritin following adjustment ranged from −0.33 to 10.07

Conclusion: BMI-based regression correction had differential effects across BMI categories and gender. The adjustment substantially attenuated higher ferritin levels in normal BMI and males. In contrast, a minimal increase was observed in overweight/obese individuals and in females with lower baseline ferritin levels. Notably, the magnitude of these percentage change in ferritin is small relative to the reported within-person biological variation (BV) of ferritin, suggesting limited clinical impact.

Highlights

- Obesity, a chronic low-grade inflammatory state, influences iron homeostasis.
- BMI-adjusted ferritin can be estimated using a regression correction approach.
- BMI-adjustment shows differential effects on serum ferritin levels in anemic individuals
- The adjustment effects varied based on gender and BMI categories
- These percentage change in ferritin observed after adjustment less likely to have a clinical impact.

Introduction

Iron is an essential trace mineral required for various biological functions such as oxygen transport and storage, electron transport chain, DNA synthesis, DNA repair, detoxification, and cofactors for certain enzymes. Levels of iron are tightly controlled and is essential for maintaining systemic and cellular iron homeostasis. Hepcidin is master regulator of systemic iron homeostasis. It binds to ferroportin and triggers its internalisation and ubiquitin mediated degradation, thereby controlling the intestinal iron absorption and release of recycled iron from macrophages that phagocytose senescent red blood cells [1].

Deficiency of iron leads to iron deficiency anemia (IDA) which is the major public health problem in India [2]. The laboratory work up of anemia includes an iron profile study which comprises of serum iron, ferritin, total iron binding capacity (TIBC), percentage transferrin saturation and complete blood count [3]. Reduced levels of serum ferritin, serum iron and percentage transferrin saturation are indicative of iron deficiency. Iron is stored as ferritin in the liver cells, and structurally a single molecule of ferritin can accommodate ~4500 iron atoms [4]. Hence, serum ferritin is considered as the maker of total body iron stores and its levels are crucial in diagnosing and monitoring iron deficiency. Iron is transported in blood by transferrin, a transporter protein produced by liver. Free iron is detrimental as it has increased propensity to form reactive oxygen species (ROS), which leads to cellular damage [5].

Ferritin is also an acute phase reactant and it is non-specifically elevated in inflammatory conditions like chronic kidney diseases, infections, autoimmune disorders, and malignancy[6]. This inflammation induced elevation of ferritin can obscure the true iron deficiency in an individual. Obesity, a growing global health concern is characterised by a state of chronic low-grade systemic inflammation. In obese adipose tissue, there is evident infiltration of macrophages leading to the production of pro-inflammatory cytokines such as interleukin-1,6 (IL-1, IL-6) and tumor necrosis factor – alpha (TNF-α) [7]. IL-6, is a potent pro-inflammatory cytokine and is known to increase the expression of ferritin and hepcidin. There are studies that has reported the influence of obesity on serum ferritin.

A study done in adolescents has shown that BMI explained the variation in serum ferritin levels to an extent that it did not reflect the true iron deficiency in those with obesity and overweight [8]. A significant linear positive association exist between serum ferritin and BMI [9] and it varies based on gender in adults [10]. Individuals with high BMI and high CRP had increased levels of serum ferritin despite low levels of iron [11]. These observations suggest serum ferritin levels tend to be higher in those with inflammation and could obscure the true iron deficiency. To address this, a study proposed calculation of inflammation adjusted serum ferritin by employing a regression-based approach to calculate CRP-adjusted ferritin. It was found a substantial variation was observed in classification of iron deficiency compared to unadjusted ferritin [12]. Studies that explored the effect of BMI-adjusted ferritin are limited.

This study employed a regression-based approach to calculate BMI-adjusted ferritin levels and to determine its effects across BMI category, gender and on classification on iron deficiency status.

Materials and Methods

Subjects

This study was done using de-identified patient data retrieved from laboratory information system (LIS) and medical records of a tertiary care hospital in South India. The data was retrieved for the period of 11 months from July 2024 to June 2025 (n=599). The data was screened and these who had complete data of serum ferritin age, gender, body weight and height were included (n=458). Individuals with missing values for any of these parameters were excluded from the analysis (n=141). The data of serum iron, total iron binding capacity and hemoglobin were retrieved. Hematological parameters were estimated on a Beckman DXH-800 analsyer. Serum iron and total iron-binding capacity were estimated by the ferrozine colorimetric method on Toshia 120 FR. Serum ferritin was estimated by electro chemiluminescence immunoassay (Roche Cobas e411).

Calculation of BMI-adjusted serum ferritin

A regression correction (RC) approach was employed to calculate BMI-adjusted serum ferritin levels [12] for the patients included in the study. A step-by-step procedure to calculate BMI-adjusted ferritin using internal regression correction is detailed in Table 1.

Table 1: Calculation of BMI-adjusted serum ferritin.

Step 1:

A: Natural log-transform (ln) the outcome variable

- ln(Ferritin_unadjusted)

(Note: BMI need not be log transformed as linear relationship exists between BMI and Ferritin)

B: Construct linear regression model (lm) to identify the beta coefficient (beta)

- lnFerritin_unadjusted = α + β × BMI
- lm (lnFerritin_unadjusted ~ BMI)

(Note: where α represents the intercept , β the regression coefficient and lm represents linear regression model)

Step 2:

A: Determine the median BMI value of the study group

The median BMI of the study population (or subgroup) was determined and used as the reference BMI (BMI_ref)

B: Calculate BMI Adjusted Ferritin using the regression correction equation

Input data from Step 1B and Step 2A.

lnFerritin_adjusted = lnFerritin_unadjusted – (β * (BMI – BMI_ref))

(Note: The adjustment was restricted to individuals whose BMI exceeded the subgroup-specific median BMI)

C: Back transform the adjusted ferritin

Adjusted Ferritin = Exp^(lnFerritin_adjusted)

The median BMI was observed to be 22.92 (males), 23.01 (females), 22.65 (normal BMI) and 27.49 (overweight/obesity). The beta coefficient for normal BMI group is 0.023511, overweight-obese group is −0.00426, males is 0.03896, females is −0.02312. The alpha-intercept for normal BMI group is 3.51348, overweight-obese group is 3.7796, males is 3.5403, females is 4.1185. These values were used to calculate BMI-based adjustments of ferritin in the respective groups . The subgroup specific equations for BMI-based adjustment of ferritin (Table 2).

Table 2: Subgroup-Specific Equations for BMI-Based Adjustment of Ferritin.

Subgroup	Equation to calculate "ln(Adjusted Ferritin)"	Condition for applying adjustment
Normal BMI	ln(Ferritin unadjusted) − 0.023511 × (BMI − 22.7)	BMI > 22.7
Overweight/Obese	ln(Ferritin unadjusted) − (−0.00426) × (BMI − 27.5)	BMI > 27.5
Males	ln(Ferritin unadjusted) − 0.03896 × (BMI − 22.9)	BMI > 22.9
Females	ln(Ferritin unadjusted) − (−0.02312) × (BMI − 23.01)	BMI > 23.01

Statistical Analysis

All statistical analysis was done using R programming language V.4.4.1. The data were checked for distribution using Shapiro Wilk Test and appropriate statistical tests were done. Chi Square test was used to compare the categorical data, and continuous data were compared using Mann-Whitney U or independent t test. Paired Wilcoxon signed rank test was done to compare the ferritin values before and after adjustment. A p-value of less than 0.05 will be considered as statistically significant.

Results

A total of 458 subjects were included in the study, comprising of 172 males (38%) and 286 females (62%). There was no difference in age and gender distribution between normal BMI group and overweight-obese group. Majority of the subjects were found to be anemic (95 % overall, 96% in normal BMI group and 95% in overweight-obese group) (Table 3). Hemoglobin, serum iron and TIBC levels were similar between the two groups. Levels of unadjusted serum ferritin were significantly lower in obese-overweight group. Levels of CRP were higher in obese-overweight group, but not statistically significant (Table 3).

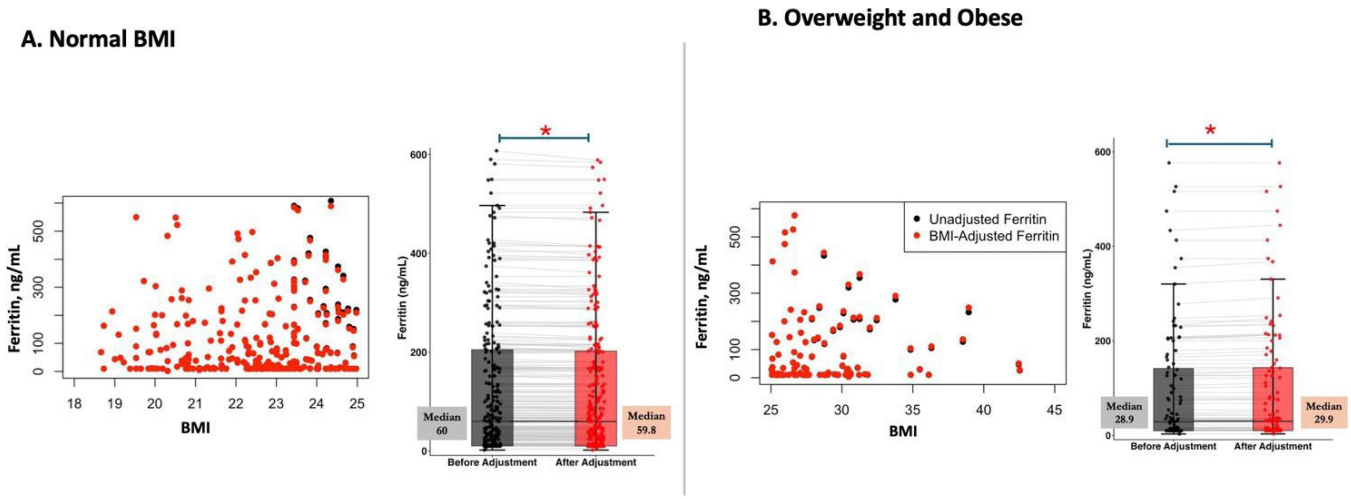
Table 3: Comparison of baseline characteristics and biochemical parameters based on BMI.

	Overall (n=458)	Normal BMI (BMI 18.5-24.9) (n=286)	Overweight and Obese (BMI >25) (n=111)	p value
Gender, n (%) Male Female	 172 (38) 286 (62)	 109 (38) 177 (62)	 36 (32) 75 (68)	0.348
Age, in years	50 (36-64)	53.5 (39-65)	49 (40-61)	0.3771
BMI, in kg/m^2	23 (21-25)	22.7 (21.4-23.4)	27.5 (26.3-30.4)	<0.0001
BMI status, n (%) Underweight Normal Overweight Obesity	 61 (13) 286 (62) 78 (17) 33 (7)	--	--	--
Anemia based on Hb* , n (%) Present Absent	 386 (95) 20 (5)	 239 (96) 10 (4)	 94 (95) 5 (5)	0.8917
Hb, g/dL$	8.5 (7-9.7)	8.4 (7.1-9.5)	8.7 (7-9.8)	0.405
Serum ferritin, ng/mL (unadjusted)	41.75 (10-179.9)	60 (10.3-205)	29 (10-142)	0.019
Serum iron, µg/dL$	27.5 (14-64)	27.5 (15-62.5)	29 (14-68)	0.800
Serum TIBC, µg/dL$	306 (210-393)	285 (202-377)	298 (241-426)	0.396
Serum CRP, mg/L$	16.9 (4.8-17.3)	17.5 (4.97-52.97)	47.65 (4.95-191)	0.113

*Data is represented in median and interquartile range. The p value is computed by comparing the obese-overweight and normal weight participants. *Anemia is diagnosed if Hb < 12g/dL in women and <13 g/dL in men. $ indicates paucity or missing data; Hb, serum iron, CRP and TIBC are available only in subset of participants.*

Regression correction was applied to derive BMI-adjusted serum ferritin values. In individuals with normal BMI, adjustment resulted in a minimal reduction in median ferritin levels (60.0 to 59.8 ng/mL; $p < 0.0001$; Figure 1A). In contrast, among obese individuals, a slight increase in median ferritin levels was observed following adjustment (28.9 to 29.9 ng/mL; $p < 0.0001$; Figure 1B).

Figure 1: Comparison of adjusted ferritin based on BMI status.



*Scatter (left) and box-and-whisker (right) plots show unadjusted (black) and BMI-adjusted (red) ferritin concentrations in normal weight group (A) and overweight-obese group (B). The adjustment was restricted to individuals whose BMI exceeded the median value of their respective group. Boxes indicate median and interquartile range and whiskers represents the values within 1.5 times the interquartile range. Grey lines denote paired observations. Paired Wilcoxon signed rank test was done to compare the ferritin values before and after adjustment. * indicates p < 0.05.*

In gender wise comparison, the age was significantly lower in females compared to males. BMI was slightly higher in females, however there were no difference in proportion of obesity, overweight between the two groups. Proportion of anemia were similar (93 % vs 96%) between two groups (Table 4). Hemoglobin and serum iron levels were similar between the two groups. Levels of unadjusted serum ferritin were significantly lower and TIBC levels were higher in females. Levels of CRP were higher in females, but not statistically significant (Table 4).

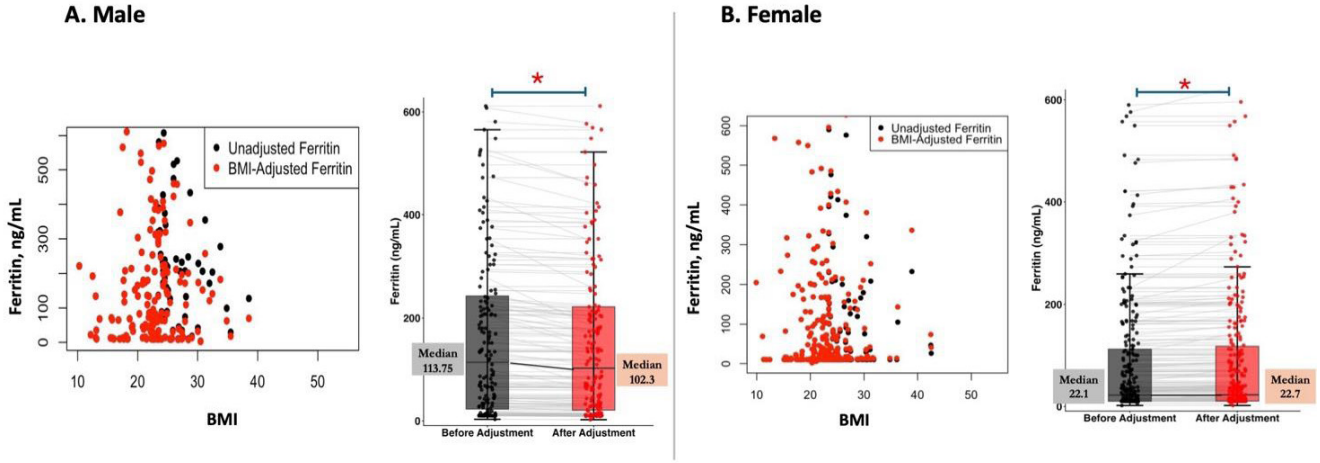
Table 4: Comparison of baseline characteristics and biochemical parameters based on gender.

	Overall (n=458)	Male (n=172)	Female (n=286)	p value
Age, in years	50 (36-64)	55 (41-68)	48 (35-60)	0.0024
BMI, in kg/m2	23 (21-25)	22.9 (20.9-24.5)	23.01 (20.8-25.1)	<0.0001
BMI status, n (%) Underweight Normal Overweight Obesity	 61 (13) 286 (62) 78 (17) 33 (7)	 27 (16) 109 (63) 25 (15) 11 (6)	 34 (12) 177 (62) 53 (19) 22 (8)	0.481
Anemia based on Hb * , n (%) Present Absent	 386 (95) 20 (5)	 140 (93) 10 (7)	 246 (96) 10 (4)	0.316
Hb, g/dL$	8.5 (7-9.7)	8.5 (7.4-9.9)	8.4 (6.8-9.5)	0.081
Serum ferritin, ng/mL (unadjusted)	41.75 (10-179.9)	113.75 (23-243)	22.1 (10-112)	<0.0001
Serum iron, µg/dL$	27.5 (14-64)	30.9 (17.8-54.5)	23 (13-70.2)	0.71
Serum TIBC, µg/dL$	306 (210-393)	237 (178-340)	343 (246-416)	0.002
Serum CRP, mg/L$	16.9 (4.8-17.3)	15.80 (4-72.8)	18.55 (5.45-97.25)	0.362

*Data is represented in median and interquartile range. The p value is computed by comparing the males and females. Anemia is diagnosed if Hb < 12g/dL in women and <13 g/dL in men. *Anemia is diagnosed if Hb < 12g/dL in women and <13 g/dL in men. $ indicates paucity or missing data; Hb, serum iron, CRP and TIBC are available only in subset of participants.*

In males, the adjustment attenuated higher ferritin values resulting in a reduction of the median ferritin level (113.75 ng/mL to 102.3 ng/mL, p < 0.0001) (Figure 2A). However, in females, the adjustment had resulted in a minimal increase in the median ferritin (22.1 ng/mL to 22.7 ng/mL, p <0.0001) (Figure 2B).

Figure 2: Gender-wise comparison of adjusted ferritin.



*Scatter (left) and box-and-whisker (right) plots show unadjusted (black) and BMI-adjusted (red) ferritin concentrations in males (A) and females (B). The adjustment was restricted to individuals whose BMI exceeded the median value of their respective group . Boxes indicate median and interquartile range and whiskers represents the values within 1.5 times the interquartile range. Grey lines denote paired observations. Paired Wilcoxon signed rank test was done to compare the ferritin values before and after adjustment. * indicates p < 0.05.*

The percentage change in ferritin before and after adjustment is shown in Table 5. It was -0.33 % in normal BMI and 3.46 % in overweight-obese group. It is -10.07 % in males and 2.71 % in females.

Table 5: Percentage change in ferritin levels before and after adjustment.

Groups	Serum Ferritin (ng/mL)		Percent change, %
	Before adjustment	After adjustment	
Normal BMI	60 (10.3-205)	59.8 (10.21-202.15)	-0.33
Overweight and Obese	28.9 (10-142)	29.9 (10.2-143.3)	3.46
Males	113.75 (23-243)	102.3 (20.9-221.8)	-10.07
Females	22.1 (10-112)	22.7 (10.3-117.8)	2.71

*Serum ferritin levels were represented in medians (interquartile range). The percent change is calculated as (Adjusted Ferritin – Unadjusted Ferritin/ Unadjusted Ferritin)*100.*

Applying a standard cut-off of 30 ng/mL to classify iron deficiency were carried out using ferritin values before and after adjustment. The proportion of iron deficiency were similar before and after adjustment in gender-wise and BMI-based comparison (Table 6).

Table 6: Classification of iron deficiency status based on unadjusted and BMI-adjusted serum ferritin.

	Before adjustment ID / Non-ID, n (%)	After adjustment ID / Non-ID, n (%)	p value
Male (n=172)	123/49 (70/30)	122/50 (71/29)	1.00
Female (n=286)	155/131 (54/46)	153/133 (53/47)	0.933
Normal BMI (n=286)	114/172 (40/60)	115/171 (40/60)	1.00
Overweight Obese BMI (n=111)	57/54 (51/49)	56/55 (50/50)	1.00

Data is represented in numbers (percentages). Ferritin <30 ng/mL: iron deficiency (ID); ferritin ≥30 ng/mL: non-iron deficiency (Non-ID). Pearson Chi-Squared test was done to compare the proportion of iron deficiency before and after adjustment in each group.

Discussion

In this study, a regression correction was applied to derive BMI-adjusted serum ferritin values in the participants, the majority of whom were anaemic. The BMI-based adjustment resulted in a minimal reduction in ferritin levels in individuals with normal BMI. However, a slight increase was observed in those with overweight and obesity. The adjustment caused substantial attenuation of higher ferritin values in males, whereas in females, the adjustment resulted in a minimal increase. There were no difference in the classification of iron deficiency status between adjusted and unadjusted ferritin levels.

Adjusting for BMI produced a minimal decline in serum ferritin levels in those normal BMI (60 vs 59.8). The median CRP levels in these patients was 17.5 mg/L (>10, cut-off for inflammation) suggesting the presence of systemic inflammation. The adjustment led to a more pronounced reduction in individuals with higher ferritin levels, which may reflect correction for inflammation-driven elevations. However, the adjustment had a very minimal reduction in those who had already low ferritin levels. This may be because these individuals are likely to be iron deficient, where the effect of iron deficiency predominates over that of adiposity-induced inflammation. The adjustment produced an opposite effect on serum ferritin levels in those with overweight and obesity i.e. adjusted ferritin levels were higher compared to unadjusted ferritin levels (28.9 vs 29.9). The levels of ferritin increased post-adjustment. This contrasts with the expectation that adiposity has a positive influence, and that BMI-based adjustment is likely to cause a reduction in serum ferritin levels. These participants also had systemic inflammation as suggested by their CRP levels (median 67.5 mg/L) indicating a substantial inflammatory burden. This highlights the intricate relationship between adiposity-associated inflammation and its impact on iron homeostasis [13]. While adiposity association inflammation induces ferritin, as a positive acute-phase reactant, it is also associated with increased expression of hepcidin [13]. Hepcidin is the master regulator of systemic iron homeostasis. It binds to and causes the degradation of ferroportin, thereby reducing intestinal iron absorption and leading to hypoferremia and iron deficiency anemia [1]. In this study, the obese overweight participants had a lower baseline ferritin value (28.9 ng/mL), so BMI-

based adjustment may produce a minimal increase in ferritin levels among individuals with low baseline ferritin, where the influence of iron deficiency outweighs that of adiposity-related inflammation. Hence, the net effect of adiposity on ferritin is not straightforward and may vary depending on baseline iron status and inflammatory state. Therefore, despite lower baseline ferritin levels in the overweight/obese group, the application of BMI-based regression correction remains physiologically justifiable, as it attempts to account for these underlying adiposity-related influences. Furthermore, the inverse relationship observed between BMI and ferritin in our dataset further supports the need to explore how such an adjustment behaves in iron-deficient populations. There is limited evidence evaluating BMI-based ferritin adjustment in populations with a high burden of anaemia.

In this study, BMI-based adjustment caused substantial attenuation of higher ferritin values in males (102.3 vs 113.75), whereas in females, the adjustment resulted in a minimal increase (22.7 vs 22.1). The unadjusted ferritin levels were higher in males despite majority of them being anemic suggesting the possible influence of inflammation rather than true iron status. This is supported by higher CRP levels in both males and females indicating an underlying inflammatory state. In contrast, lower baseline ferritin levels in females are more consistent with iron deficiency, potentially explaining the minimal increase after adjustment. This pattern is also observed in those with overweight and obesity, where the influence of iron deficiency may predominate alongside coexisting inflammation. Gender difference in the levels of ferritin are well known and may be attributed to varied reasons such as hormonal influences, dietary iron intake and menstrual blood loss [14,15].

BMI-based regression correction resulted in modest and variable changes in ferritin across subgroups, with a minimal reduction in individuals with normal BMI (−0.33%), slight increases in those with overweight/obesity (+3.10%) and females (+2.71%), and a more pronounced reduction in males (−10.07%). Importantly, the magnitude of these changes is small relative to the reported within-person biological variation (BV) of ferritin (5.37%–39.13%, mean 21.64%) [16] , suggesting that most differences are unlikely to be clinically meaningful.

Adjusted ferritin did not impact the classification of iron deficiency. This could be because most of the study participants were anemic (i.e. 95%) with median Hb of 8.5 g/dL. Additionally, the serum iron levels were low in the participants, indicating iron deficiency as the major possibility of anemia.

Limitations

It is a retrospective study and included data retrieved from medical records and LIS. This study design is not ideal to determine the causal effect of BMI on iron deficiency and serum ferritin. In this study, there is no data regarding the presence of other inflammatory conditions, dietary history and intake of iron supplements. The sample size is limited and there is paucity of data on iron related proteins and inflammatory marker. A larger sample size with equally distributed non-anemic participants would be needed to confirm these findings.

Conclusion

BMI-based regression correction had differential effects across BMI categories and gender. The adjustment substantially attenuated higher ferritin levels in normal BMI and males. In contrast, a very minimal increase was observed in overweight/obese individuals and in females with lower baseline ferritin levels. Notably, the magnitude of these percentage change in ferritin is small relative to the reported within-person biological variation (BV) of ferritin, suggesting limited clinical impact. These findings may be explained by the predominance of iron deficiency over adiposity-related inflammatory effects in individuals with lower baseline ferritin.

Implications

In the Indian context, where obesity and iron deficiency anemia are increasingly prevalent, BMI-adjusted ferritin may offer additional value in interpreting ferritin variability and in more accurately identifying true iron deficiency.

Author's Disclosures

Conflict of interests

The authors do not have any competing interests to declare in this study.

Ethics approval and consent to participate

This study is approved by Institutional Ethics Committee (VMCIEC/184/2025).

Authors contribution

JR: Study Design, Conceptualization, Writing – Original Draft, Reviewing And Editing, Data Visualization, Validation, Methodology, Formal Analysis And Supervision. DB – Formal analysis, Data curation, Writing – review and editing

Funding

This study does not require any funding.

Availability of data

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

Not applicable.

References

1. Nemeth E, Tuttle MS, Powelson J, Vaughn MB, Donovan A, Ward DM, et al. Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. Science 2004;306:2090–2093 DOI: https://doi.org/10.1126/science.1104742
2. Anemia Mukt Bharat. Anemia Mukt Bharat 2025https://anemiamuktbharat.info/anemia-situation/india (accessed May 30, 2025)
3. Rifai N, Horvath AR, Wittwer CT. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics (8th ed.). Elsevier. vol. 2. 8 th. 2023
4. Kenelly PJ, Murray RK, Jacob M, Varghese J. Chapter 52. Plasma Proteins & Immunoglobulins. Harpers Illus. Biochem. 31st ed., LANGE - McGraw Hill Education; 2018
5. Oh HL, Lee JA, Kim DH, Lim JS. Reference values for serum ferritin and percentage of transferrin saturation in Korean children and adolescents. Blood Res 2018;53:18–24 DOI: https://doi.org/10.5045/br.2018.53.1.18
6. Ems T, St Lucia K, Huecker MR. Biochemistry, Iron Absorption. StatPearls, Treasure Island (FL): StatPearls Publishing; 2024
7. Heslin AM, O'Donnell A, Buffini M, Nugent AP, Walton J, Flynn A, et al. Risk of Iron Overload in Obesity and Implications in Metabolic Health. Nutrients 2021;13:1539 DOI: https://doi.org/10.3390/nu13051539
8. Shattnawi KK, Alomari MA, Al-Sheyab N, Bani Salameh A. The relationship between plasma ferritin levels and body mass index among adolescents. Sci Rep 2018;8:15307 DOI: https://doi.org/10.1038/s41598-018-33534-4
9. Han H, Ni P, Zhang S, Ji X, Zhu M, Ma W, et al. The association of body mass index and weight waist adjustment index with serum ferritin in a national study of US adults. Eur J Med Res 2023;28:374 DOI: https://doi.org/10.1186/s40001-023-01343-9
10. Han L, Wang Y, Li J, Zhang X, Bian C, Wang H, et al. Gender differences in associations of serum ferritin and diabetes, metabolic syndrome, and obesity in the China Health and Nutrition Survey. Mol Nutr Food Res 2014;58:2189–2195 DOI: https://doi.org/10.1002/mnfr.201400088
11. Alam F, Memon AS, Fatima SS. Increased Body Mass Index may lead to Hyperferritinemia Irrespective of Body Iron Stores. Pak J Med Sci 2015;31:1526 DOI: https://doi.org/10.12669/pjms.316.7724
12. Namaste SM, Rohner F, Huang J, Bhushan NL, Flores-Ayala R, Kupka R, et al. Adjusting ferritin concentrations for inflammation: Biomarkers Reflecting Inflammation and Nutritional Determinants of Anemia (BRINDA) project. Am J Clin Nutr 2017;106:359S-371S DOI: https://doi.org/10.3945/ajcn.116.141762
13. Ramasamy J. The Dysregulation of Iron Homeostasis in Obesity. In: Preedy VR, Patel VB, editors. Handb. Public Health Nutr. Int. Natl. Reg. Perspect., Cham: Springer Nature Switzerland; 2025, p. 1–21 DOI: https://doi.org/10.1007/978-3-031-32047-7_115-1
14. Bachman E, Feng R, Travison T, Li M, Olbina G, Ostland V, et al. Testosterone suppresses hepcidin in men: a potential mechanism for testosterone-induced erythrocytosis. J Clin Endocrinol Metab 2010;95:4743–4747 DOI: https://doi.org/10.1210/jc.2010-0864
15. Skolmowska D, Głąbska D, Skolmowska D, Głąbska D. Analysis of Heme and Non-Heme Iron Intake and Iron Dietary Sources in Adolescent Menstruating Females in a National Polish Sample. Nutrients 2019;11 DOI: https://doi.org/10.3390/nu11051049
16. Gnanou JV, Muthayya S, Kurpad AV. Biological variation of plasma ferritin in healthy adult males in south Indian population—A sample study. Indian J Clin Biochem 2006;21:193–195 DOI: https://doi.org/10.1007/BF02913095

Research Article

Analytical Performance of an Alternative Wavelength in the O-CPC Calcium Assay

Serdar Hira[1*], Ömer Özcan[1], Osman Metin İpcioğlu[1], Mustafa Gültepe[1]

[1]*Department of Clinical Biochemistry, GATA Haydarpaşa Training and Research Hospital, İstanbul, Türkiye*

Article Info

[*]***Corresponding Author:***
Serdar Hira
Department of Clinical Biochemistry, Royal Hospital
Sağlık Street No:4, 10200, Paşakonak, Bandırma, Balıkesir, Türkiye
E-mail: mdserdarhira@gmail.com
Telephone: +90 545 543 39 69

Keywords

Analytical performance, Calcium, Clinical laboratory, Method comparison, O-cresolphthalein complexone, Wavelength selection

Abstract

Objective: The o-cresolphthalein complexone (O-CPC) method is widely used for serum calcium measurement but may exhibit analytical limitations related to background absorbance and bias, particularly at low calcium concentrations. This study evaluated the analytical performance of an alternative wavelength (524 nm) compared with the conventional 572 nm measurement.

Methods: Spectral analysis of the calcium–O-CPC complex identified a secondary absorbance feature near 536 nm. Because this wavelength was not available in routine analysers, 524 nm was selected as the nearest operational alternative for analytical evaluation. Performance was assessed using multipoint calibration models, control sera, precision analysis, and comparison with an automated analyzer in 40 samples, including prepared mixtures to extend the analytical range. Agreement between methods was evaluated using regression and Bland–Altman analyses.

Results: Reagent blank absorbance was lower at 524 nm compared with 572 nm (0.048 vs 0.122). Linearity was high at both wavelengths ($R^2 \approx 0.998$). Precision analysis demonstrated comparable CV% values between wavelengths. Regression analysis showed a slope closer to unity and a smaller intercept at 524 nm relative to 572 nm. Bland–Altman analysis demonstrated comparable overall agreement between methods at both wavelengths. In the low concentration subgroup (<9 mg/dL), mean negative bias was reduced from −0.55 mg/dL at 572 nm to −0.30 mg/dL at 524 nm.

Conclusions: Measurement at 524 nm demonstrated comparable overall analytical performance to the conventional 572 nm wavelength and was associated with reduced negative bias in the low concentration range. These findings suggest that alternative wavelength selection may represent a practical approach for optimizing analytical performance in routine O-CPC calcium assays.

Introduction

Accurate measurement of serum calcium is essential in clinical laboratory practice and plays a central role in the evaluation of metabolic, endocrine, and renal disorders [1]. Analytical reliability is particularly important near clinical decision thresholds, where small analytical deviations may influence the classification of hypocalcemia and subsequent clinical interpretation.

In routine laboratories, calcium is commonly measured using colorimetric methods based on metallochromic indicators [2]. Among these, the o-cresolphthalein complexone (O-CPC) method is widely used due to its simplicity, cost-effectiveness, and compatibility with automated systems. In alkaline conditions, O-CPC forms a colored calcium complex typically measured at wavelengths around 570–580 nm, corresponding to the primary absorbance maximum [3].

Despite its widespread use, the O-CPC method may exhibit analytical limitations, particularly in the low concentration range [3]. Background absorbance and signal-to-noise characteristics may contribute to negative bias near clinically relevant decision thresholds [4]. While reagent composition and calibration strategies have been extensively investigated, the analytical impact of wavelength selection has received comparatively less attention in routine clinical applications [4,5].

Wavelength selection is generally treated as a fixed parameter in automated assays; however, it may influence analytical performance by affecting background signal and bias distribution [4]. Identification of alternative measurement regions within the absorbance spectrum may therefore provide a practical approach for evaluating analytical performance without modifications of reagent chemistry or instrumentation [6].

The aim of this study was to evaluate the analytical performance of an alternative wavelength in the O-CPC calcium assay, with particular emphasis on the low concentration range relevant to clinical decision thresholds.

Materials and Methods

Study design and instrumentation

This analytical method comparison study was conducted in the clinical biochemistry laboratory of a tertiary care hospital. Absorbance spectra of the calcium–O-CPC complex were obtained using a spectrophotometer (Philips AU 530). Spectral analysis revealed a secondary absorbance feature near 536 nm in addition to the primary peak at 572 nm. Because 536 nm was not operationally available in routine clinical analysers, 524 nm was selected as the nearest available alternative wavelength for analytical evaluation. The objective was not to maximize absorbance intensity, but to evaluate whether an alternative spectral region with lower background absorbance could influence analytical performance. Measurements performed at 524 nm were compared with the conventional 572 nm wavelength.

Routine measurements were compared with an automated analyzer (Architect c8000, Abbott Laboratories, USA) operating under manufacturer-recommended analytical conditions.

Calibration and linearity assessment

Linearity was evaluated by calibrators at concentrations of 0, 4, 6, 8, 10, and 12 mg/dL. Two multipoint calibration models (0–6–10 mg/dL and 0–4–8 mg/dL) were applied. Linearity was assessed using linear regression analysis and expressed as the coefficient of determination (R^2).

Agreement assessment and precision

Agreement with assigned concentrations was evaluated using two levels of commercial control sera (Bio-Rad Liquicheck, Germany) with assigned values of 8.64 mg/dL and 11.9 mg/dL. Measured values were compared with assigned concentrations and expressed as percentage bias (%). Precision was assessed by repeated measurements and expressed as coefficient of variation (CV%).

Sample selection and preparation

A total of 40 samples were included in the analysis. Twenty routine patient serum samples were analyzed. To extend the analytical measurement range, an additional 20 prepared samples were generated using predefined mixing protocols. Dilution samples were prepared by mixing serum with distilled water at ratios of 1:1 and 2:1 (serum:water), while higher concentration samples were generated by mixing serum with a calcium calibrator (13 mg/dL) at ratios of 1:3 and 1:4. These prepared mixtures were included to ensure distribution across the analytical range; however, potential matrix-related effects were considered during interpretation of the findings.

Method comparison

Spectrophotometric measurements at 524 nm and 572 nm were compared with results obtained from the automated analyzer. Automated measurements were performed in duplicate and averaged.

Agreement between methods was evaluated using linear regression, Passing–Bablok regression, Deming regression, and Bland–Altman analyses.

Statistical analysis

Statistical analyses were performed using SPSS version 24 (IBM Corp., USA). Correlation between methods was assessed using Pearson correlation coefficient (r). Mean bias was calculated as the average difference between methods. Limits of agreement were calculated as mean difference ±1.96 standard deviations. Subgroup analysis was performed for samples with calcium concentrations <9 mg/dL. Method comparison analysis was performed in alignment with CLSI EP09-A3 principles for analytical method comparison. Regression parameters were additionally reported with 95% confidence intervals where applicable.

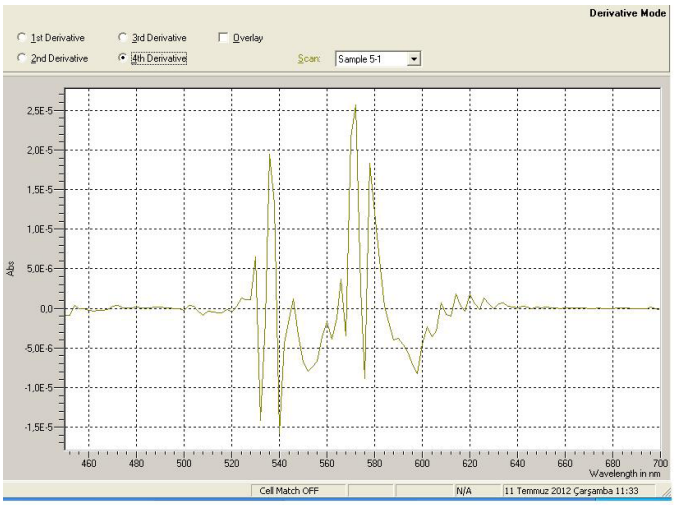
Ethical considerations

The study was performed in the clinical biochemistry laboratory of GATA Haydarpaşa Training and Research Hospital, İstanbul, Turkey. The study used anonymized residual serum samples obtained during routine laboratory practice. According to institutional policy applicable at the time of the study, formal ethics committee approval and informed consent were not required for analytical evaluations performed on anonymized residual samples.

Results

Spectral analysis demonstrated a dominant absorbance peak at 572 nm and a secondary spectral feature near 536 nm (Figure 1). Based on this observation, subsequent analytical evaluations were performed at 524 nm as the nearest operationally available alternative wavelength.

Figure 1: Derivative Spectral Analysis of the Calcium–O-CPC Complex.



Fourth-order derivative spectrum demonstrating the primary absorbance peak at 572 nm and a secondary spectral feature near 536 nm. The secondary feature was used to guide selection of an alternative wavelength for analytical evaluation.

Table 1: Summary of Analytical Performance at 524 nm and 572 nm.

Parameter	524 nm	572 nm
Reagent blank absorbance	0.048	0.122
Linearity (R^2)	0.998	0.998
Regression slope	1.03	1.05
Intercept (mg/dL)	−0.16	−0.53
Correlation (r)	0.965	0.969
Mean bias (mg/dL)	+0.125	−0.082
Bias <9 mg/dL (mg/dL)	−0.30	−0.55

R^2: coefficient of determination; r: Pearson correlation coefficient; bias: mean difference relative to automated analyzer measurements.

Linearity analysis demonstrated excellent linearity at both wavelengths across calibration models (Table 1). Using the 0–6–10 mg/dL calibration model, R^2 values were 0.999 at 524 nm and 0.998 at 572 nm. Similar results were obtained with the 0–4–8 mg/dL calibration model (R^2 = 0.998 at 524 nm and R^2 = 0.996 at 572 nm).

Precision analysis demonstrated comparable repeatability at both wavelengths. At 524 nm, CV values were approximately 2.6% for the low-level control and 2.9% for the high-level control. Corresponding CV values at 572 nm were approximately 2.2% and 3.5%, respectively (Table 2).

Control serum measurements demonstrated numerically lower bias at 524 nm for the low-level control material (Table 2). For the low-level control (8.64 mg/dL), measured values were 8.43 mg/dL at 524 nm and 7.98 mg/dL at 572 nm. For the high-level control (11.9 mg/dL), measured values were 12.08 mg/dL at 524 nm and 11.79 mg/dL at 572 nm. Agreement analysis for the conventional wavelength is presented in Supplementary Figure S1.

Table 2: Control Serum Agreement and Precision Analysis.

Control Level	Target (mg/dL)	524 nm (mg/dL)	572 nm (mg/dL)	Bias 524 (%)	Bias 572 (%)	CV% 524	CV% 572
Low	8.64	8.43	7.98	−2.4%	−7.6%	2.6%	2.2%
High	11.9	12.08	11.79	+1.5%	−0.9%	2.9%	3.5%

Bias (%) was calculated relative to assigned control values. CV: coefficient of variation.

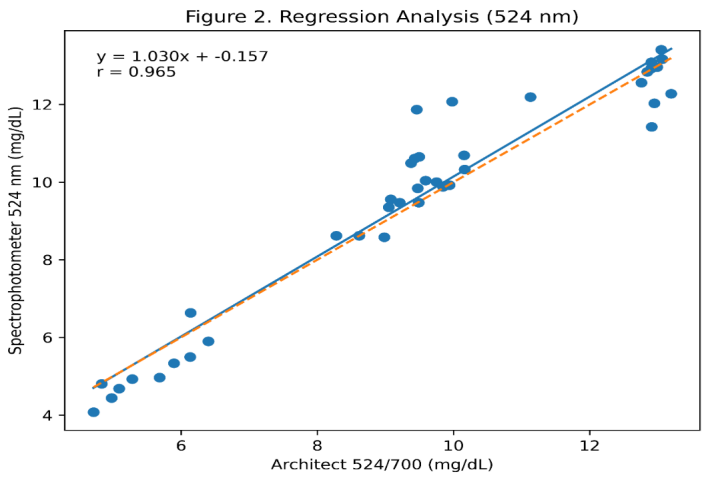
Method comparison analysis in 40 samples demonstrated strong correlation between spectrophotometric and automated measurements at both wavelengths (r = 0.965 at 524 nm; r = 0.969 at 572 nm). Regression analysis demonstrated a slope numerically closer to unity at 524 nm compared with 572 nm (1.03 vs 1.05), together with a smaller intercept magnitude (−0.16 vs −0.53 mg/dL) (Table 3, Figure 2). The corresponding 95% confidence intervals are presented in Table 3. Additional regression analyses using Passing–Bablok and Deming methods yielded comparable results (Supplementary Table S1).

Table 3: Regression and Agreement Analysis Parameters (n = 40).

Parameter	524 nm	572 nm
Slope (95% CI)	1.03 (0.94-1.12)	1.05 (0.96-1.13)
Intercept (mg/dL, 95% CI)	−0.16 (-1.05 to 0.74)	−0.53 (-1.38 to 0.32)
Correlation (r)	0.965	0.969
Standard Error of Estimate	0.045	0.043
Mean Bias (mg/dL)	+0.125	−0.082
Lower Limits of Agreement (mg/dL)	−1.38	−1.55
Upper Limits of Agreement (mg/dL)	+1.63	+1.38

CI: confidence interval; limits of agreement calculated using Bland–Altman analysis.

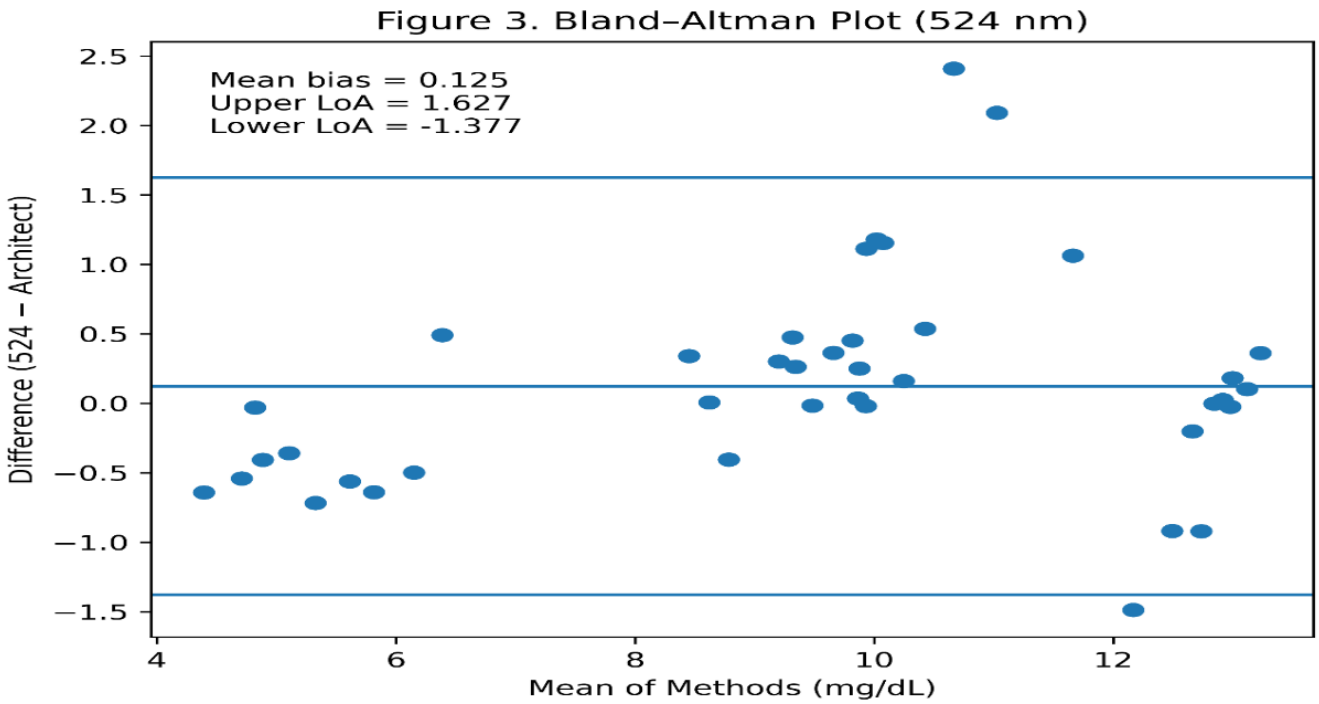
Figure 2: Regression Analysis between Spectrophotometric Measurement at 524 nm and Architect Autoanalyzer Results.



Linear regression analysis comparing spectrophotometric measurements performed at 524 nm with Architect autoanalyzer results in 40 samples. The solid line represents the regression line, and the dashed line represents the identity line (y = x).

Bland–Altman analysis demonstrated comparable overall agreement between methods at both wavelengths, with a mean bias of +0.125 mg/dL at 524 nm and −0.082 mg/dL at 572 nm (Figure 3).

Figure 3: Bland–Altman Analysis of Measurements at 524 nm.



Bland–Altman plot showing mean bias and limits of agreement between spectrophotometric measurements performed at 524 nm and Architect autoanalyzer results in 40 samples.

In the low concentration subgroup (<9 mg/dL), mean negative bias was lower at 524 nm compared with 572 nm (−0.30 vs −0.55 mg/dL) (Table 1).

No statistically significant concentration-dependent bias was observed for either wavelength.

Discussion

In this study, we evaluated the analytical performance of an alternative wavelength in the O-CPC calcium assay under routine laboratory conditions. Measurements performed at 524 nm demonstrated comparable overall agreement with the conventional 572 nm wavelength, together with lower negative bias in the low concentration range.

Spectral analysis identified a secondary absorbance feature near 536 nm in addition to the primary peak at 572 nm. Because this wavelength is not operationally available in routine analysers, 524 nm was selected as the nearest practical alternative for analytical evaluation. The objective was not to maximize absorbance intensity, but to assess whether an alternative spectral region with lower background absorbance could influence analytical performance characteristics. Previous studies evaluating derivative spectrophotometric approaches and wavelength optimization strategies have similarly demonstrated that alternative spectral regions may influence analytical signal characteristics and background interference [4,6].

One of the most consistent findings was the substantial reduction in reagent blank absorbance at 524 nm. Lower background absorbance may improve signal-to-background characteristics, particularly in measurements near the lower analytical range [7]. In the subgroup analysis, negative bias in samples with calcium concentrations below 9 mg/dL was lower at 524 nm compared with 572 nm. Precision profiles were comparable between wavelengths, suggesting that this observation was not solely attributable to analytical imprecision.

Although the absolute difference in bias was modest, small analytical deviations near clinical decision thresholds may influence classification of borderline hypocalcemia. In this context, reduction of negative bias in the low concentration range may still be analytically relevant, particularly when considered within acceptable total allowable error limits for calcium measurement [5].

The calcium–O-CPC reaction may demonstrate concentration-dependent nonlinearity related to complex stoichiometry at low calcium concentrations and reagent limitation at higher concentrations, supporting the routine use of two-point or multipoint calibration strategies [8]. In this context, the consistent linearity observed at 524 nm across both 0–6–10 and 0–4–8 calibration schemes suggests stable analytical response characteristics across the evaluated measurement range.

Regression analyses demonstrated a slope numerically closer to unity and a smaller intercept magnitude at 524 nm relative to 572 nm. However, confidence intervals overlapped between methods, indicating broadly comparable overall analytical agreement.

Importantly, the present findings do not suggest replacing the conventional wavelength solely on the basis of correlation equivalence. Instead, they demonstrate that wavelength selection represents a controllable analytical parameter that can be adjusted to improve performance in specific concentration ranges. Because this approach does not require modification of reagent composition or instrumentation, it may represent a practical strategy for analytical optimization in routine laboratory settings.

From a clinical laboratory perspective, even modest differences in analytical performance may have meaningful implications when applied across large testing volumes. In particular, reduced negative bias in the low concentration range may support more consistent classification near hypocalcemia decision thresholds. The primary contribution of this study is not the introduction of a new analytical method, but the demonstration that wavelength selection may influence analytical characteristics within clinically relevant concentration ranges.

This study has several limitations. The automated analyzer was used as a comparative method rather than a definitive reference method such as isotope dilution mass spectrometry or atomic absorption spectroscopy [9]. Second, the sample size was modest and the study should therefore be considered an exploratory analytical evaluation rather than a full-scale clinical validation study. Third, prepared mixture samples were included to extend the analytical measurement range; however, dilution-related matrix effects may have influenced analytical behavior and limit direct comparability with native patient samples [10,11]. In addition, potential analytical interference from hemolysis, lipemia, bilirubin, or magnesium was not specifically evaluated. Because wavelength availability differs among analytical platforms, the findings may be partially instrument-dependent and require validation in other automated systems before broader generalization. Finally, because the study was performed using a single analytical platform, external validation on other systems is required.

In conclusion, measurement at 524 nm demonstrated comparable overall analytical performance to the conventional 572 nm wavelength, together with lower negative bias in the low concentration range. These findings suggest that alternative wavelength selection may represent a practical and cost-neutral approach for analytical optimization of routine O-CPC calcium assays.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical approval

Ethical approval and informed consent were waived according to institutional policy applicable at the time of the study because anonymized residual samples obtained during routine laboratory practice were used.

Author contributions

S.H. designed the study and wrote the manuscript. Ö.Ö. contributed to data analysis. O.M.İ. and M.G. supervised the study and reviewed the manuscript. All authors approved the final version.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

1. Torfadóttir JE, Uusi-Rasi K. Calcium – a scoping review for Nordic Nutrition Recommendations 2023. Food Nutr Res 2023;67:10303. https://doi.org/10.29219/fnr.v67.10303
2. Verkhratsky A, Petersen OH, editors. Calcium Measurement Methods. New York City, United States: Humana Press; 2010.
3. Burtis CA, Bruns DE, editors. Tietz textbook of clinical chemistry and molecular diagnostics. 6th ed. St Louis: Elsevier; 2018.
4. Frans SD, Harris JM. Selection of analytical wavelengths for multicomponent spectrophotometric determinations. Anal Chem 1985;57:2680–2684. https://doi.org/10.1021/ac00290a055
5. Jassam N, Thomas A, Hayden K, Dearman R, Barth JH. The impact of the analytical performance specifications of calcium and albumin on adjusted calcium. Annals of Clinical Biochemistry. 2020;57(5):382-388. https://doi.org/10.1177/0004563220944426
6. El-Sayed AA, El-Salem NA. Recent developments of derivative spectrophotometry and their analytical applications. Anal Sci 2005;21:595–614. https://doi.org/10.2116/analsci.21.595
7. Chandarana C, Juwarwala I, Shinde R. UV derivative spectroscopy: a comprehensive review for advancement in quantification. Pharm Chem J 2025;58:1755–1773. https://doi.org/10.1007/s11094-025-03335-x
8. Corns CM, Ludman CJ. Some observations on the nature of the calcium–cresolphthalein complexone reaction and its relevance to the clinical laboratory. Ann Clin Biochem 1987;24:345–351. https://doi.org/10.1177/000456328702400403
9. Han B, Ge M, Zhao H, Yan Y, Zeng J, Zhang T, et al. Determination of serum calcium levels by 42Ca isotope dilution inductively coupled plasma mass spectrometry. Clin Chem Lab Med 2018;56:51–58. https://doi.

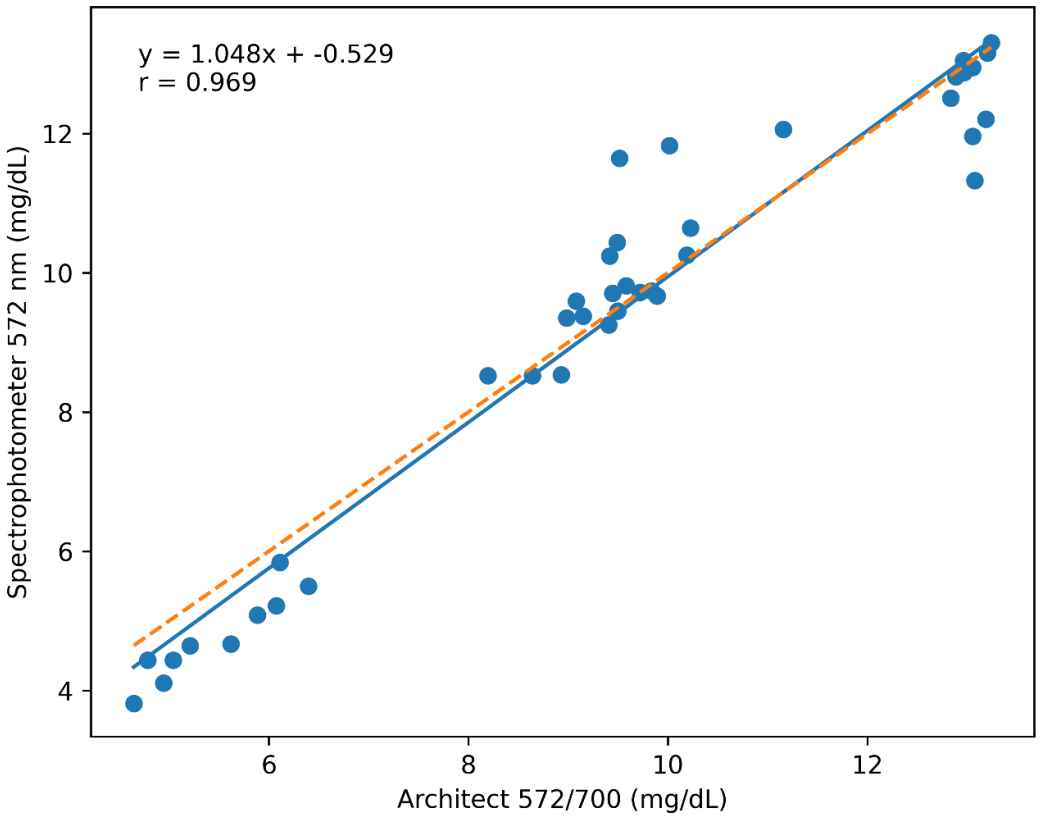
org/10.1515/cclm-2017-0175
10. Guideline IH. Bioanalytical method validation and study sample analysis M10. ICH Harmonised Guideline: Geneva, Switzerland. 2022. Accessed: 02/04/2026
11. Pum J. A practical guide to validation and verification of analytical methods in the clinical laboratory. Advances in clinical chemistry. 2019;90:215-281. https://doi.org/10.1016/bs.acc.2019.01.006

Supplementary Table S1: Agreement analysis using alternative regression models.

Method	Slope	Intercept
Linear regression	0.98	0.14
Passing–Bablok	0.98	0.20
Deming	0.98	0.10

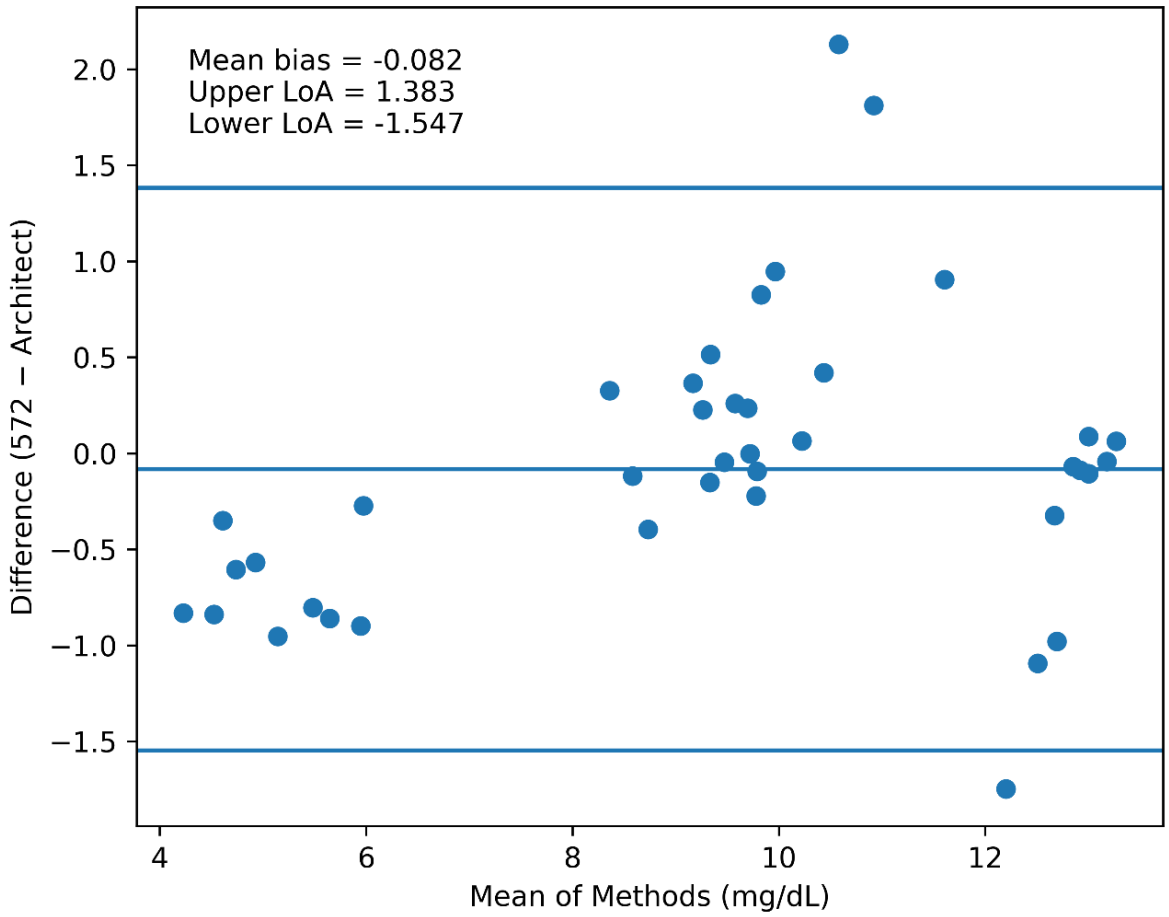
Comparison of regression parameters obtained using different analytical models (linear regression, Passing–Bablok, and Deming regression) for spectrophotometric measurements at 524 nm. Slope and intercept values are presented. Results are presented for comparison purposes and reflect model-specific estimates obtained using different regression approaches. Minor differences in slope and intercept values are expected due to methodological differences between regression models.

Supplementary Figure 1: Regression Analysis between Spectrophotometric Measurement at 572 nm and Architect 572/700 nm Method.



Linear regression analysis showing agreement between spectrophotometric measurements performed at the conventional 572 nm wavelength and Architect autoanalyzer results (572/700 nm). The solid line represents the regression line, and the dashed line represents the identity line.

Supplementary Figure 2: Bland–Altman Agreement Analysis for 572 nm Measurement.



Bland–Altman plot demonstrating mean bias and limits of agreement between spectrophotometric measurements at 572 nm and Architect 572/700 nm results.

Research Article

Impact of Hemolysis on UCH-L1 Quantification in Serum and Plasma

Paula García Martín[1*], Silvia Álvarez Rodríguez[1], Covadonga Solar Macarro[1], Rafael Venta Obaya[1] Eduardo Martínez-Morillo[1]

[1]*Department of Biochemistry San Agustin University Hospital, Camino de Heros, Avilés, Asturias, Spain*

Article Info

[*]***Corresponding Author:***
Paula García Martín
Department of Biochemistry, San Agustin University Hospital,
Camino de Heros, 6, 33401, Avilés, Asturias, Spain
E-mail: paula.garciammamar@sespa.es

Keywords

Hemolysis, plasma, serum, traumatic brain injury, ubiquitin carboxy-terminal hydrolase L1

Abstract

Introduction: Ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) is a neuron-enriched deubiquitinase widely used for the evaluation of mild traumatic brain injury (TBI). Although minimal hemolysis-related interference has been reported for the Abbott TBI assay, routine laboratory observations indicate that hemolysis may substantially influence UCH-L1 measurements. This study aimed to evaluate the effect of hemolysis on UCH-L1 concentrations in plasma and serum samples and its potential clinical implications.

Methods: Hemolysis interference was evaluated in accordance with CLSI C56-A and EP07 guidelines using pooled plasma samples and paired individual patient samples. Controlled hemolysis was induced by adding graded amounts of erythrocyte hemolysate to non-hemolyzed samples. UCH-L1 concentrations were measured using the Abbott Alinity i system. In addition, a retrospective analysis of routine TBI biomarker requests was conducted to estimate the potential rate of hemolysis-induced false-positive results.

Results: In pooled plasma, UCH-L1 increased linearly (R^2 = 0.992) with the hemolysis index (HI), with a 10% analytical bias observed at HI values of 100 or lower. Erythrocyte-derived samples showed UCH-L1-related signals more than 20-fold higher than plasma. In individual patient samples, hemolysis induced a dose-dependent but highly variable positive bias. Retrospective analysis suggested that hemolysis could result in false-positive UCH-L1 results in up to 21.5% of TBI-negative patients with concentrations near the clinical cut-off.

Conclusions: Hemolysis induces a clinically relevant positive bias in UCH-L1 measurements, even at moderate HI levels. These findings underscore the importance of strict preanalytical control and cautious interpretation of UCH-L1 results obtained from hemolyzed samples.

Abbreviations: CT: Computed Tomography; CV: Coefficient of variation; GFAP: Glial Fibrillary Acidic Protein; HI: Hemolysis Index; I: Interfered; Int (%): Percentage of interference; NI: Non-interfered; TBI: Traumatic Brain Injury; UCH-L1: Ubiquitin Carboxy-Terminal Hydrolase L1; UCH-L3: Ubiquitin Carboxy-Terminal Hydrolase L3.

Introduction

Ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) is a member of the ubiquitin C-terminal hydrolase family of deubiquitinases, which also includes UCH-L3, UCH-L5, and BRCA1 Associated Protein 1. Its primary function is to hydrolyze ubiquitin from small C-terminal extensions, and it is highly expressed in neurons, with lower expression in gonadal tissue [1,2]. Due to its neuronal enrichment, UCH-L1 was identified in the early 2000s as a promising biomarker for traumatic brain injury (TBI). In combination with glial fibrillary acidic protein (GFAP), UCH-L1 is currently used in clinical practice to aid in the evaluation of patients with suspected mild TBI (Glasgow Coma Scale score of 13–15). Large validation studies have demonstrated that this biomarker panel can reduce the number of unnecessary cranial computed tomography (CT) scans by approximately 40%, thereby limiting radiation exposure and long-term risks associated with imaging [3].
Abbott's TBI assay is an immunoassay panel for in vitro diagnostics used to measure GFAP and UCH-L1 in human plasma and serum, with established cut-offs of 35 and 400 pg/mL for GFAP and UCH-L1, respectively. This assay aids in the evaluation of patients aged over 18 years old with suspected mild TBI within 12 hours of injury. The results help determine whether a cranial CT scan is required. A result exceeding either threshold is considered positive and supports the need for CT evaluation [4].
According to the Abbott package insert (H22846R01, B4Z233, created in 2021) for TBI assay, hemolysis does not produce a significant interference (≤10% bias) in UCH-L1 measurement at hemoglobin concentrations up to 1,000 mg/dL, which are generally considered equivalent to a hemolysis index (HI) of approximately 1,000 arbitrary units [4]. Despite this manufacturer's claim, our laboratory observations suggest that hemolysis may influence measured UCH-L1 concentrations. To further investigate this potential source of analytical variability, the present study was designed to assess the effect of hemolysis on plasma and serum UCH-L1 concentrations using both pooled and individual samples.

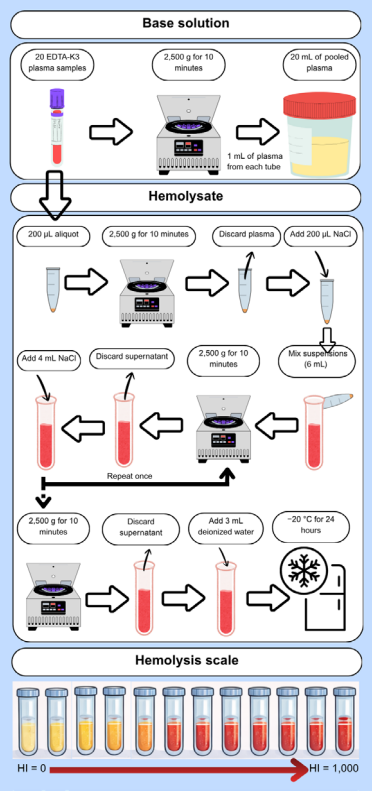
Material and Methods

The procedure for the evaluation of endogenous interferences was adapted from the CLSI C56-A and EP07 guidelines [5,6], with minor modifications.

Pooled sample preparation

A total of 20 EDTA-K3 plasma samples (3 mL purple-top tubes, VACUETTE®) from primary care patients undergoing routine complete blood count testing (presumably healthy individuals) were collected within 24 h of venipuncture, prior to disposal. Only samples visually assessed as non-hemolyzed and with volumes close to the maximum tube capacity (3 mL) were selected. As samples were anonymized immediately after collection by removing all identifying labels prior to specimen handling, written informed consent was not required. This study was approved by the Research Ethics Committee of the Principality of Asturias (approval number: 2025.386).
The UCH-L1 assay is manufacturer-validated for serum and plasma (EDTA-K2/K3 and lithium heparin) [4]. EDTA-K3 plasma was selected for the pooled experimental phase for practical reasons: only samples no longer required for clinical care were deliberately chosen for research use, and hemolysate preparation requires whole blood prior to centrifugation. Unlike serum and lithium heparin samples, which are processed immediately after collection with plasma already separated from the cellular fraction, EDTA tubes are the only collection format routinely stored as whole blood.
Each of the 20 tubes was inverted 10 times to ensure homogenization. Subsequently, to obtain the base solution, the samples were centrifuged at 2,500 g for 10 minutes. From each tube, 1 mL of plasma was collected and pooled into a sterile container, resulting in a total volume of 20 mL of pooled plasma, which served as the base solution for all subsequent experiments (Figure 1).

Figure 1: Schematic representation of pooled plasma preparation and the development of the hemolysis scale.



To obtain the hemolysate, a 200 µL aliquot of whole blood was collected from each EDTA-K3 tube and transferred to a 1.5 mL Eppendorf tube. Following centrifugation (2,500 g, 10 min), the plasma (100 µL, approximately) was discarded, and 200 µL of isotonic saline (NaCl 154 mmol/L) was added to each cellular pellet. The twenty resulting suspensions (300 µL, approximately) were gently mixed and subsequently pooled into a 15 mL Falcon® conical tube, yielding a final volume of approximately 6 mL. This procedure minimized the risk of hemolysis due to ABO incompatibility. Then, the pooled cell suspension was centrifuged (2,500 g, 10 min). The supernatant was discarded, and the pellet was washed twice with 4 mL of isotonic saline solution. After each wash, the suspension was centrifuged again (2,500 g, 10 min) and the supernatant was completely removed. Subsequently, 3 mL of deionized water was added to the washed pellet to induce cell lysis, and to achieve a HI of 10,000 arbitrary units. The resulting lysate was homogenized and stored at −20 °C for 24 hours. After thawing, the hemolysate was thoroughly mixed and centrifuged at 2,500 g for 30 minutes to remove cellular debris (Figure 1).

Two sample types were prepared using the base solution and the hemolysate:

Non-interfered (NI) sample, consisting of 9 mL of base solution mixed with 1 mL of deionized water.

Interfered (I) sample, consisting of 9 mL of base solution mixed with 1 mL of the hemolysate (HI = 10,000).

Twelve samples were prepared by mixing defined volumes of the NI and I base solutions. Serial combinations of these two matrices yielded samples with estimated HI values ranging from 0 (no hemolysis) to approximately 1,000.

To assess assay precision in the presence of hemolysis, a pooled sample was prepared from 7 EDTA-K3 plasma specimens using the standardized hemolysis protocol, targeting HI of approximately 0, 250, 500, and 1000. HI and UCH-L1 concentrations were measured in quintuplicate for each sample, with imprecision expressed as the coefficient of variation (CV). UCH-L1 concentrations were determined using a chemiluminescent microparticle immunoassay on the Abbott Alinity i module, whereas HI were measured on the Alinity c module using spectrophotometric detection of free hemoglobin, with absorbance readings taken at four wavelength pairs (500/524, 572/604, 628/660, and 524/804 nm) [7].

The degree of interference (Int (%)) was expressed as the percentage difference between the I and NI samples, calculated as Int (%) = 100 * (Ci - C0) / C0, where Ci and C0 represent the measured concentrations in each I and NI samples, respectively. A maximum allowable error of 10% was used as the criterion to define the acceptable limit of interference.

Selection and processing of individual hemolyzed samples

According to our standard operating procedure for sample rejection, a new blood sample is requested whenever the HI measured in plasma or serum exceeds 500. For this study, 20 samples were randomly selected–15 lithium heparin plasma

samples and 5 serum samples–with HI values above 500, for which a second, non-hemolyzed sample was obtained within one hour of the initial sample collection.

Unlike the pooled experimental phase, lithium heparin plasma and serum were used in this phase, as these are the most processed matrices in our laboratory. This allowed hemolysis interference to be assessed under real-world conditions using clinically rejected, hemolyzed samples, each paired with a non-hemolyzed sample from the same patient.

After routine analyses were completed, 700 µL of each sample (600 µL for mixture preparation and 100 µL to account for dead volume) was transferred into 6 mL polystyrene tubes. Only samples from patients without requests for brain trauma markers were included, to avoid potential physiological increases in these markers during the interval between sample collections. Additionally, samples were excluded if less than 500 µL of non-hemolyzed serum or plasma remained available for potential further analyses. All samples were fully anonymized prior to analysis.

Paired non-hemolyzed and hemolyzed samples from each patient were combined to produce four graded levels of hemolysis. The volume of non-hemolyzed serum or plasma was progressively decreased from 300 µL to 0 µL, while the volume of the corresponding hemolyzed sample was increased from 0 µL to 300 µL, resulting in mixtures with hemolyzed fractions of approximately 0%, 33%, 67%, and 100%. UCH-L1 concentrations and HI were measured in each mixture, and the Int (%) was calculated.

To verify the manufacturer's claim of assay suitability across serum, EDTA plasma and lithium heparin plasma, a matrix comparison study was performed. Blood samples were collected from three healthy volunteers in all three tube types, and the standardized hemolysis protocol was applied to generate both non-hemolyzed (HI ≈ 0) and hemolyzed (HI ≈ 1,000) conditions. UCH-L1 concentrations and HI values were measured in triplicate, and differences among matrices were evaluated by repeated-measures analysis of variance (ANOVA) using MedCalc software.

Analysis of Results from Patients with a Routine Request for TBI Markers

Data from patients for whom the UCH-L1 test was requested at our hospital between August 2024, and July 2025 were retrospectively retrieved from the laboratory information system. The dataset was analyzed to assess the potential impact of hemolysis on the occurrence of false-positive UCH-L1 results and its possible implications for clinical decision-making in TBI.

Only samples corresponding to requests for TBI biomarkers were included. To minimize the influence of severe preanalytical interference, analyses were restricted to samples with a HI below 100. No clinical records were reviewed, and no personal or identifiable patient data were accessed or collected.

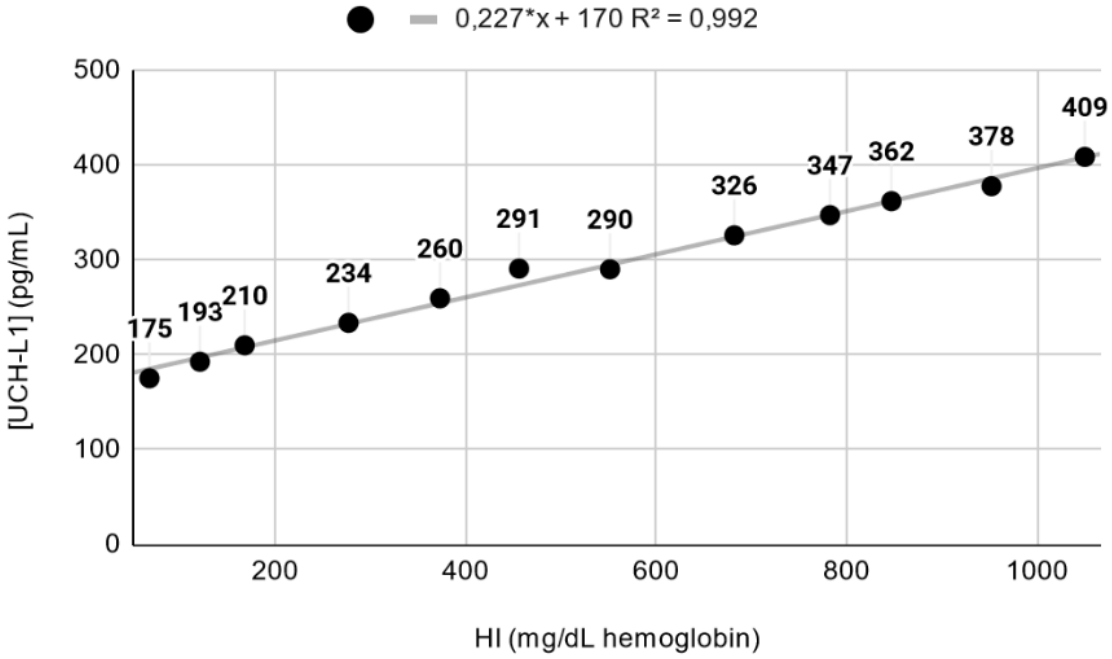
Results

Pooled samples

In the pooled plasma experiment evaluating imprecision, the mean HI values across quintuplicates were 27, 240, 447, and 867, with corresponding mean UCH-L1 concentrations of 196, 305, 399, and 556 ng/L. CVs for UCH-L1 were 6.1%, 8.9%, 5.1%, and 8.6%, respectively, showing no apparent association with the degree of hemolysis, while CVs for HI remained below 3% in all cases. Fitting a quadratic model to the mean values, a 10% increase in UCH-L1 concentration was estimated at HI = 63, and an absolute concentration of 400 ng/L, corresponding to the manufacturer's recommended decision threshold for ruling out brain injury, was estimated at HI = 525. Nearly identical estimates were obtained when the model was fitted using only the first individual measurement from each quintuplicate (HI = 63 and 523, respectively).

In the 20 pooled plasma experiment, UCH-L1 concentrations increased linearly with increasing HI (Figure 2). The regression equation was UCH-L1 = 0.227 (95% confidence interval: 0.212-0.242) x HI + 170 (95% confidence interval: 161-179); with an R^2 value of 0.992. The baseline UCH-L1 concentration at HI = 68 was 175.4 pg/mL. According to this model, a 10% deviation from baseline (corresponding to an increase of approximately 17 pg/mL) would be reached at an HI of 101.

Figure 2: Linear regression of UCH-L1 (pg/mL) concentration versus hemolysis index (HI) in pooled plasma. The line represents the best-fit linear model, illustrating the dose-dependent increase of UCH-L1 with increasing hemolysis.



Assuming a linear increase in UCH-L1 concentration of 0.227 pg/mL per HI unit, we evaluated which HI values would result in UCH-L1 concentrations exceeding 400 pg/mL, depending on the initial baseline concentration. Samples with baseline concentrations below 175 pg/mL did not exceed 400 pg/mL within the HI range tested; samples with baseline concentrations between 175 and 300 pg/mL exceeded 400 pg/mL at intermediate HI values (approximately 400–1,000); and samples with baseline concentrations above 300 pg/mL consistently exceeded 400 pg/mL when HI values were greater than 440 (Table 1).

Table 1: Estimated hemolysis index (HI) required for a sample to exceed the clinical ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) threshold of 400 pg/mL, based on a linear increase in UCH-L1 concentration with hemolysis (0.227 × HI).

Baseline UCH-L1 concentration (pg/mL)	Required increase in UCH-L1 (pg/mL)	Required HI (mg/dL hemoglobin)
150	250	1,101
175	225	991
200	200	881
250	150	661
300	100	441
350	50	220
375	25	110

For each baseline UCH-L1 concentration, the corresponding HI value required to reach 400 pg/mL was calculated. Samples with baseline UCH-L1 concentrations below 175 pg/mL required very high HI values (> 950) to exceed the decision threshold, whereas samples with baseline concentrations of ≥ 300 pg/mL surpassed 400 pg/mL at moderate levels of hemolysis (HI < 450). The calculated increase in UCH-L1 above baseline (shown in the middle column) represents the additional concentration required to reach the pathological cut-off.

The relative abundance of UCH-L1 —or the interfering substance detected by the assay— was also evaluated within erythrocytes compared with plasma. To generate a hemolysate, 2 mL of washed red blood cells were mixed with 3 mL of deionized water, resulting in a 5 mL preparation (I sample) with an approximate hematocrit of 40%, comparable to that of an average individual. The UCH-L1 concentration measured in this erythrocyte-derived matrix was 408.8 pg/mL (Figure 2). In contrast, the mean UCH-L1 concentration in non-hemolyzed patient plasma was approximately 175 pg/mL, slightly overestimated due to a baseline HI of 68 rather than 0. Based on these measurements, the intracellular-to-plasma ratio was approximately 23.5, indicating that the erythrocyte compartment contains more than 20-fold higher UCH-L1–related signal than plasma.

Individual samples

In individual patient samples, UCH-L1 concentrations increased progressively with increasing HI, confirming a positive, dose-dependent interference pattern. This trend was consistent with that observed in the pooled plasma experiment, although considerable variability among patients was evident. To better characterize this relationship, individual quadratic regression models were generated for each patient (Table 2). The quadratic model provided a superior fit compared with a simple linear model, reflecting the non-linear nature of hemolysis interference.

Table 2: Calculated hemolysis index (HI), mg/dL hemoglobin, required to increase ubiquitin carboxy-terminal hydrolase L1 (UCH-L1) concentration by 10% and to reach 400 pg/mL, based on quadratic models.

Patient	Baseline UCH-L1 concentration (pg/mL)	Baseline HI	UCH-L1 concentration increased by 10%	HI required for 10% UCH-L1 increase	HI required to reach 400 pg/mL of UCH-L1	Quadratic equation
1	44	3	48	17	2106	43,5 + 0,259x - 4,26E-05x² =y
2*	58	0	64	13	856	58.7 + 0.405x - 7.17E−06x² =y
3*	65	11	71	40	2275	64,6 + 0,161x - 5,97E-06x² =y
4	68	1	74	13	754	66.9 + 0.549x - 1.42E−04x² =y
5	81	2	89	42	1795	78,3 + 0,257x - 7,16E-05x² =y
6*	96	1	106	39	1929	95,5 + 0,272x - 7,05E-05x² =y
7	106	1	116	26	723	105 + 0.441x - 4.59E−05x² =y
8	122	4	134	24	559	122 + 0,502x - 7,68E-06x² =y
9*	134	1	148	194	2103	134 + 0.0668x + 2.84E−05x² =y
10*	136	1	149	73	1484	139 + 0.134x + 2.82E−05x² =y
11	140	3	154	25	369	134 + 0.807x - 2.33E−04x² =y
12	142	10	156	42	775	139 + 0,41x - 9,47E-05x² =y
13	158	5	174	84	1310	156 + 0,216x - 2,27E-05x² =y
14	169	59	186	102	1009	153 + 0,341x - 1,69E-04x²=y
15	178	2	196	33	573	177 + 0.583x - 3.38E−04x²=y
16	232	110	255	171	468	171 + 0,492x - 5,63E-06x²=y
17	247	38	272	195	968	241 + 0.158x + 6.44E−06 x²=y
18	262	73	288	138	269	154 + 1,03x + -4,29E-04x²=y
19	413	54	455	95	N/A	350 + 1,13x + -2,02E-04x²=y
20	1183	164	1301	808	N/A	1206 + -0,232x + 4,33E-04x²=y

The equations were derived from individually fitted quadratic models describing the relationship between the HI and the measured UCH-L1 concentration in samples with progressively increasing hemolysis. The patients' results are shown in order of increasing baseline UCH-L1 concentration.

**Serum samples (*); remaining samples are lithium heparin plasma.*

N/A: not appicable

Baseline UCH-L1 concentrations among patients ranged from 44 to 413 pg/mL, with one patient exhibiting an exceptionally high concentration of 1,183 pg/mL. Calculated UCH-L1 concentrations corresponding to a 10% deviation from baseline varied from 48 to 455 pg/mL, with an outlier sample of 1,301 pg/mL (patient previously mentioned). The HI required to produce a 10% change in UCH-L1 concentration varied between 13 and 195 for most patients, except for one patient, for whom it was 808. Finally, the HI necessary to raise UCH-L1 above the cut-off value of 400 pg/mL (resulting in a false-positive) ranged from 269 to 1,000 in 10 of the 20 patients (50% potential false-positives), while it exceeded 1,000 in the remaining patients, except for patients 19 and 20, whose baseline UCH-L1 concentration was already over 400 pg/ml.

In the matrix comparison study evaluating assay suitability across serum, EDTA plasma, and lithium heparin plasma, the mean HI was 8 in non-hemolyzed samples and 1144 in hemolyzed samples. Under non-hemolyzed conditions, mean UCH-L1 concentrations for the three volunteers were 44.5, 69.8, and 101 ng/L, rising to 308.1, 318.0, and 451.8 ng/L under hemolyzed conditions. No statistically significant differences among matrices were observed under either condition ($p > 0.05$).

Analysis of Results from Patients with a Routine Request for TBI Markers

During the study period, 986 patients presenting with mild TBI underwent assessment, and TBI markers were requested. Among them, 260 patients showed concentrations of both markers (GFAP and UCH-L1) below their respective cut-off values (35 and 400 pg/mL, respectively), with 56 of these "TBI-negative" patients exhibiting UCH-L1 concentrations in the range of 300–400 pg/mL. Based on our regression model, all these patients would have had UCH-L1 concentrations above 400 pg/mL if their sample HI had exceeded 440. This indicates that analyzing hemolyzed samples (HI > 440) from these patients would have resulted in a 21.5% false-positive rate among all patients with a negative TBI result (56 of 260).

Discussion

The linear relationship observed between UCH-L1 concentration and HI demonstrates that increasing hemolysis contributes to higher UCH-L1 values, which can substantially affect the clinical interpretation of this biomarker in plasma and serum. In pooled plasma, the baseline UCH-L1 concentration was 175.4 pg/mL, in close agreement with previous reports indicating that basal levels in patients with a negative TBI are generally below 200 pg/mL [4,8,9] using Alinity ci. This consistency supports the validity of our pooled model. Sample matrix (serum, EDTA plasma or lithium heparin plasma) does not appear to be a determining factor in hemolysis-induced interference, and the use of single measurements rather than replicates does not meaningfully affect the estimates obtained, as these remained consistent with those derived from the quintuplicate determinations.

According to the relationship [UCH-L1] ≈ 0.227 × HI, the degree of hemolysis required for UCH-L1 concentrations to exceed the clinical threshold of 400 pg/mL was evaluated. For baseline UCH-L1 concentrations below 175 pg/mL, only severe hemolysis (HI > 950) could result in falsely elevated values above the threshold; however, samples with such high hemolysis would be rejected. In contrast, when the baseline UCH-L1 concentration exceeded 300 pg/mL, hemolysis levels of approximately 450 were sufficient to raise the calculated concentration above 400 pg/mL (Table 1). These results indicate that the susceptibility to hemolysis-induced false-positives depends strongly on both baseline UCH-L1 concentration and the degree of hemolysis.

Applying the manufacturer's recommended allowable error of 10%, an interference threshold at HI ≈ 100 (corresponding to UCH-L1 = 193 pg/mL) was identified. Interestingly, the manufacturer's insert reports no interference from hemoglobin up to 1,000 mg/dL [4]. In contrast, our findings indicate that beyond HI ≈ 100, hemolysis can meaningfully affect UCH-L1 measurement accuracy. This observation is supported by the FDA memorandum on the Banyan Brain Trauma Indicator, which reported that hemoglobin produced statistically significant interference beyond the prespecified acceptance limit of 500 mg/dL, while showing no interference at concentrations up to 62.5 mg/dL [10].

From a clinical perspective, these findings are relevant because UCH-L1 is increasingly used to guide imaging decisions in TBI. In our laboratory, the clinical decision limit for recommending CT imaging is 400 pg/mL [4,11]. According to the regression model (UCH-L1 = 0.227 x HI + 170), an HI greater than approximately 1,013 would be necessary to raise a baseline UCH-L1 concentration above this pathological threshold, potentially resulting in false-positive interpretations and unnecessary imaging. At such extreme levels of hemolysis, samples should be rejected and recollected; however, lower HI values may be problematic when other published thresholds are applied. For example, the decision cut-offs proposed by Papa et al. (2022), Karamian et al. (2025), and Bazarian et al. (2018, 2021)—189, 225, 327, and 360 pg/mL, respectively [10,12–14]—would correspond to HI values of approximately 84, 242, 692, and 837 according to our model. This suggests that even moderate hemolysis could artifactually elevate UCH-L1 concentrations above these clinical cut-offs, potentially generating false-positive results. It should be noted that these cut-offs were derived using ELISA-based assays in all studies except for Bazarian et al. (2021), who employed the i-STAT Alinity system [14], which has been seen to be highly correlated to Alinity ci [15]. However, these studies differ not only in their proposed clinical cut-offs but also in their analytical methodologies. As shown by the 2024 SEQC-ML external quality assessment for serum indices, different

analysers yield non-identical HI values even when measuring equivalent samples [16]. Such heterogeneity limits the direct comparability of absolute UCH-L1 values and interference thresholds across studies. Nevertheless, regardless of the analytical approach used, our findings highlight that hemolysis can shift UCH-L1 values toward commonly used decision limits, potentially leading to misclassification. This may not only increase healthcare costs but also expose patients to unnecessary radiation and delay appropriate clinical decision-making.

This study also demonstrates that the effect of hemolysis on UCH-L1 measurements is highly variable among patients, and that individual hemolysis–response curves can differ substantially even under identical analytical conditions. Baseline UCH-L1 concentrations ranged from 44 to 413 pg/mL in most individuals, with one outlier reaching approximately 1,183 pg/mL (Table 2). Patients with low baseline UCH-L1 concentrations showed the greatest susceptibility to hemolysis. In these cases, small increases in HI were sufficient to exceed the 10% error threshold, making their results particularly vulnerable to clinical distortion. This supports a proportional relationship in which lower initial UCH-L1 concentrations require smaller increments in HI to generate unacceptable analytical error.

The observed increase in UCH-L1 concentration with rising hemolysis may be explained by several, non-mutually exclusive mechanisms. Immunoassays might be particularly susceptible to hemolysis-related interference, as red blood cell–derived components, including free hemoglobin and intracellular proteases, may degrade target antigens or disrupt antibody–antigen binding, thereby altering signal generation [17]. However, the interference observed may not be directly attributable to free hemoglobin itself, as the manufacturer reports no significant effect at concentrations up to 1,000 mg/dL, but rather to other components released from intraerythrocytic content during hemolysis, which is consistent with the reported impact of non-specific analytical factors such as fibrin, erythrocyte remnants, platelets, or other suspended particles on assay performance [4]. Alternatively, hemolysis may lead to a true release of UCH-L1 or UCH-L1–like proteins from blood cells. Given that UCH-L3 shares approximately 52% sequence homology with UCH-L1, cross-reactivity was considered as a potential contributor [2]; however, the Banyan Brain Trauma Indicator decision memorandum [10], as well as the Abbott package insert [4], reported no detectable cross-reactivity between UCH-L3 and the UCH-L1 assay, making immunological cross-detection an unlikely explanation for the observed bias. Regardless of the underlying mechanism, our data clearly demonstrate that hemolysis is a major pre-analytical variable for UCH-L1 quantification.

Overall, the patient-level analyses indicate that hemolysis interference is neither linear nor uniform and is influenced by both baseline UCH-L1 concentrations and the specific curvature of each HI–UCH-L1 relationship. These findings emphasize that UCH-L1 results from hemolyzed samples must be interpreted with caution, as the magnitude of interference can range from negligible to clinically significant and may bias diagnostic decision-making. The data analysis highlights that even moderate hemolysis can shift UCH-L1 concentrations toward the decision cut-off and generate false-positive results, potentially affecting patient classification in diagnostic algorithms.

To our knowledge, this is the first study to describe hemolysis interference in UCH-L1 measurement at such a low HI. According to Puravet A. et al., UCH-L1 exhibits a positive bias exceeding 10% at HI of 400, identifying this value as analytically relevant [18]. However, their study did not evaluate lower hemolysis ranges due to technical limitations in HI measurement. In contrast, our analysis specifically investigated lower HI values using patient-derived samples and demonstrates that a ≥10% positive bias can already occur at HI ≈ 100, indicating that clinically meaningful interference arises well before the previously reported limit.

Limitations

Limitations include the small number of pooled and individual samples and the focus on a single analytical method. Further studies should evaluate larger, more heterogeneous cohorts, and aim to establish consensus interference thresholds for UCH-L1.

Conclusion

We have demonstrated that hemolysis produces a strong, dose-dependent positive bias in UCH-L1 concentrations, with even moderate hemolysis capable of altering values sufficiently to affect clinical interpretation, particularly at lower thresholds. The linear relationship observed in pooled plasma, together with the high erythrocyte enrichment of UCH-L1 (or the interfering substance) and the potentially substantial proportion of false-positive results in patient samples attributable to hemolysis, underscores the biomarker's marked vulnerability to red blood cell lysis. Consequently, unrecognized hemolysis may lead to false-positive classifications and unnecessary CT imaging in patients with mild TBI. These findings highlight the need for systematic hemolysis detection, assay-specific interference thresholds, and stringent sample quality criteria to ensure the reliable clinical application of UCH-L1.

Research ethics

The study was conducted in accordance with the Declaration of Helsinki. Ethical approval was granted by the Research Ethics Committee of the Principality of Asturias (September 2025, ID: 2025.386)

Informed consent

Not applicable.

Author contributions

All authors have accepted responsibility for the entire content

of this manuscript and approved its submission.

Paula García Martín: Investigation, Validation, Formal analysis, Writing - Original Draft, Visualization.
Silvia Álvarez Rodríguez: Validation, Writing - Review & Editing.
Covadonga Solar Macarro: Validation, Writing - Review & Editing.
Rafael Venta Obaya: Writing - Review & Editing.
Eduardo Martínez Morillo: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Writing - Review & Editing, Supervision.

Use of Large Language Models, AI and Machine Learning Tools
None declared.

Conflict of interest
The authors state no conflict of interest.

Research funding
None declared.

Data availability
Not applicable.

References
1. Diaz-Arrastia R, Wang KKW, Papa L, Sorani MD, Yue JK, Puccio AM, et al. Acute biomarkers of traumatic brain injury: relationship between plasma levels of ubiquitin C-terminal hydrolase-L1 and glial fibrillary acidic protein. J Neurotrauma. 2014;31(1):19–25. doi: 10.1089/neu.2013.3040.
2. Bishop P, Rocca D, Henley JM. Ubiquitin C-terminal hydrolase L1 (UCH-L1): structure, distribution and roles in brain function and dysfunction. Biochem J. 2016;473(16):2453–2462. doi: 10.1042/BCJ20160082.
3. Kobeissy F, Arja RD, Munoz JC, Shear DA, Gilsdorf J, Zhu J, et al. The game changer: UCH-L1 and GFAP-based blood test as the first marketed in vitro diagnostic test for mild traumatic brain injury. Expert Rev Mol Diagn. 2024;24(1–2):67–77. doi: 10.1080/14737159.2024.2306876.
4. Abbott Laboratories. Insert TBI: GFAP, UCH-L1 for use with Alinity i (04Z23, H22846R01, B4Z233). 2021.
5. EP07. Interference Testing in Clinical Chemistry; 2018. Available from: https://clsi.org/shop/standards/ep07/ [Accessed 23 Sep 2025].
6. C56. Hemolysis, Icterus, and Lipemia/Turbidity Indices as Indicators of Interference in Clinical Laboratory Analysis; 2012. [cited 2025 Sep 23]. Available from: https://clsi.org/shop/standards/c56/ [Accessed 23 Sep 2025].
7. Abbott. Alinity ci-series - Operations Manual 80000071-113. 2025. p. 479.
8. Lapić I, Rogić D, Bingula M, Miklić L, Gornik I. Age-Adjusted Cut-Off Values for Glial Fibrillary Acidic Protein and Ubiquitin Carboxy-Terminal Hydrolase L1 Improve the Diagnostic Accuracy of the Abbott Mild Traumatic Brain Injury Assay. Diagnostics. 2025;15(9):1093. doi: 10.3390/diagnostics15091093.
9. Ladang A, Vavoulis G, Trifonidi I, Calluy E, Karagianni K, Mitropoulos A, et al. Increased specificity of the "GFAP/UCH-L1" mTBI rule-out test by age dependent cut-offs. Clin Chem Lab Med CCLM. 2025;63(5):995–1003. doi: 10.1515/cclm-2024-1034.
10. Evaluation of automatic class III designation for Banyan Brain Trauma Indicator. Decision memorandum. Available from: https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN170045.pdf? [Accessed 03 Oct 2025].
11. FDA. Substantial equivalence determination decision summary. Available from: https://www.accessdata.fda.gov/cdrh_docs/reviews/K223602.pdf? [Accessed 03 Oct 2025].
12. Papa L, Ladde JG, O'Brien JF, Thundiyil JG, Tesar J, Leech S, et al. Evaluation of Glial and Neuronal Blood Biomarkers Compared With Clinical Decision Rules in Assessing the Need for Computed Tomography in Patients With Mild Traumatic Brain Injury. JAMA Netw Open. 2022;5(3):e221302. doi: 10.1001/jamanetworkopen.2022.1302.
13. Karamian A, Farzaneh H, Khoshnoodi M, Maleki N, Rohatgi S, Ford JN, et al. Accuracy of GFAP and UCH-L1 in predicting brain abnormalities on CT scans after mild traumatic brain injury: a systematic review and meta-analysis. Eur J Trauma Emerg Surg Off Publ Eur Trauma Soc. 2025;51(1):68. doi: 10.1007/s00068-024-02697-3.
14. Bazarian JJ, Welch RD, Caudle K, Jeffrey CA, Chen JY, Chandran R, et al. Accuracy of a rapid glial fibrillary acidic protein/ubiquitin carboxyl terminal hydrolase L1 test for the prediction of intracranial injuries on head computed tomography after mild traumatic brain injury. Acad Emerg Med. 2021;28(11):1308–1317. doi: 10.1111/acem.14366.
15. Oris C, Khatib-Chahidi C, Pereira B, Bailly Defrance V, Bouvier D, Sapin V. Comparison of GFAP and UCH-L1 Measurements Using Two Automated Immunoassays (i-STAT® and Alinity®) for the Management of Patients with Mild Traumatic Brain Injury: Preliminary Results from a French Single-Center Approach. Int J Mol Sci. 2024;25(8):4539. doi: 10.3390/ijms25084539.
16. Bullich S, Caballero A, Comas I, García C, Gómez-Rioja R, Llopis MA, et al. VII External Quality Assurance Program for Serum Indices of the Spanish Society of Laboratory Medicine (2024) [Spanish]. Annual report 2024. Accesed from: https://www.contcal.org/qcweb/

Documents/90%20Avaluacio%20anual/170%20Programas%202024/05%20%C3%8Dndices%20s%C3%A9ricos%202024.pdf
17. Simundic AM, Baird G, Cadamuro J, Costelloe SJ, Lippi G. Managing hemolyzed samples in clinical laboratories. Crit Rev Clin Lab Sci. 2020;57(1):1–21. doi: 10.1080/10408363.2019.1664391.
18. Puravet A, Sapin V, Bouillon-Minois JB, Pereira B, Durif J, Bouthors B, et al. Performance of GFAP and UCH-L1 compared to S100B in detecting intracranial injury: influence of age, hemolysis, neurodegenerative diseases, and extracranial fractures in a prospective cohort of over 1,000 patients. Clin Chem Lab Med CCLM. 2025;64(2):368–378. doi: 10.1515/cclm-2025-1129.

Research Article

Using Google Trends to Predict Dengue Outbreaks in Pakistan: A Time-Series and Causality Analysis

Alizeh Sonia Fatimi[1], Tazeen Fatima[2], Asghar Nasir[3], Muhammad Shariq Shaikh[4], Imran Siddiqui[5], Sibtain Ahmed[5*]

[1]*Medical College, Aga Khan University, Karachi, Pakistan*
[2]*Pathology and laboratory services, NICVD Karachi*
[3]*Section of Molecular Pathology, Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*
[4]*Section of Hematology, Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*
[5]*Section of Chemical Pathology, Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*

Article Info

[*]***Corresponding Author:***
Sibtain Ahmed
Assistant Professor & Consultant Chemical Pathologist
Department of Pathology and Laboratory Me
Aga Khan University Hospital
Stadium Road Karachi Pakistan
Phone: +92 21-34861951
E-mail: sibtain.ahmed@aku.edu

Abstract

Background: Dengue fever is a significant public health issue in Pakistan. Seasonal outbreaks lead to considerable illness and economic strain. Traditional surveillance systems, like the Integrated Disease Surveillance and Response (IDSR) system, depend mainly on passive reporting. They often identify outbreaks only after transmission has already worsened. Digital data sources, such as Google Trends, might provide early signs of increasing disease activity by reflecting changes in public information-seeking behavior.

Objective: To determine if Google search activity for dengue-related terms can predict dengue case counts in Pakistan during 2024.

Methods: Weekly dengue case count data were collected from the IDSR Weekly Bulletin. Weekly Relative Search Volume (RSV) data for "dengue," "dengue fever," "dengue symptoms," and "dengue treatment" were gathered from Google Trends. Case counts were adjusted to a 0–100 scale. Pearson correlation was used to examine the relationship between search activity and case count data. Time-series causality analysis was carried out using the Toda–Yamamoto method, following tests for stationarity and the specification of the Vector Autoregression (VAR) model.

Results: Strong positive correlations were found between dengue-related search interest and case count data ($r = 0.756–0.832$; $p < 0.001$). The strongest correlation was for the term "dengue." Toda–Yamamoto causality testing showed that search activity significantly predicted subsequent increases in dengue case counts for "dengue symptoms" ($p = 0.004$), "dengue treatment" ($p = 0.001$), and "dengue" ($p = 0.004$), with an approximate two-week lag.

Conclusion: Google Trends search data effectively reflects and can anticipate dengue case count patterns in Pakistan. This suggests that it could serve as a cost-effective, real-time early warning tool to complement the national surveillance system. Combining digital search monitoring with IDSR reporting could lead to earlier detection and a more proactive public health response.

Introduction

Dengue fever is one of the most common mosquito-borne infections in Pakistan. It causes thousands of cases each year and continues to challenge the country's public health system [1]. The Dengue virus (DENV), a member of the Flavivirus genus, has four antigenically distinct serotypes (DENV-1–4) spreads through Aedes aegypti and Aedes albopictus mosquitoes, which breed in stagnant water during and after the monsoon season [2]. Seasonal dengue outbreaks happen almost every year, with major peaks usually occurring between September and November. Most cases come from the Punjab and Sindh provinces, especially in large urban centers like Karachi, Lahore, and Rawalpindi [3]. According to national surveillance data, more than 28,427 confirmed cases and multiple fatalities were reported in 2024 [4]. This made it one of Pakistan's more severe outbreaks in recent years.

Several factors contribute to the increasing frequency and intensity of dengue outbreaks in the country. These factors include unplanned urban growth, inadequate drainage, and sanitation systems, and shifting rainfall patterns due to climate change [3]. A study conducted in four major Pakistani cities estimated the average direct cost per dengue patient to be Rs. 35,823 (approximately US $358 at the time of the study), highlighting the substantial economic burden associated with the disease. This cost reflects hospital admissions, diagnostic testing, vector control measures, and loss of productivity [5]. Even though dengue outbreaks are predictable, early warning systems are still weak. Outbreaks are frequently noticed only after hospitals report a rise in patients [6].

The Integrated Disease Surveillance and Response (IDSR) system is Pakistan's main framework for tracking and reporting dengue cases. It gathers weekly outbreak reports from both public and private healthcare facilities. However, the system mainly uses passive surveillance [7]. This approach can cause delays in reporting and lead to an undercount of actual cases. Many areas also have limited laboratory confirmation, making it hard to produce timely and accurate outbreak alerts. Furthermore, IDSR has been described as inconsistent and lacking in resources [8]. There is significant variation in how it is implemented across provinces. The system suffers from a shortage of trained epidemiological staff, poor coordination between provincial and federal authorities, and limited financial resources [8]. These issues weaken routine data collection and timely response efforts. As a result, by the time data are compiled and analyzed, community transmission is often already widespread. This reduces chances for early intervention [8].

To address these gaps, several countries have used digital tools for early outbreak detection. Google Trends is one such platform that gives real-time data on how often people search for specific terms [9]. Research from countries like India and Indonesia has shown that increases in dengue-related searches often happen before official numbers rise (9–12). For instance, studies in India found that online searches for "dengue symptoms" increased around 4 weeks before confirmed outbreaks [12]. This suggests that how people search for information may signal early signs of disease spread.

In Pakistan, internet access has grown quickly over the last ten years, with approximately 27% of the population now using the internet [13]. This digital data could offer useful insights into public health trends. However, there is little research on whether Google search activity can help predict dengue outbreaks locally.

This study aims to investigate the connection between Google Trends data and reported dengue case counts in Pakistan during 2024. By looking at weekly search interest for dengue-related terms and comparing it with official data from the IDSR system, this research examines if online search behavior can act as an early warning sign for upcoming outbreaks.

Methods

Data Sources

Weekly data on dengue case counts in Pakistan for 2024 came from the IDSR Weekly Bulletin. This bulletin is published by the National Institutes of Health (NIH) Pakistan through its Public Health Bulletin portal (https://phb.nih.org.pk/integratedisease-surveillance-and-response). The IDSR bulletin gathers weekly reports from provincial and district health departments throughout Pakistan, covering both public and private healthcare facilities. These reports include the counts of laboratory-confirmed and suspected dengue cases from all participating districts primarily using NS1 antigen, ELISA, IgM/IgG serology, and RT-PCR where available. However, laboratory confirmation coverage remains uneven across provinces, which can delay accurate case reporting. A substantial proportion of dengue infections are mildly symptomatic, which means true transmission may be even higher than reported.

Weekly Google Trends data were gathered from the Google Trends platform (https://trends.google.com) for the search terms "dengue fever," "dengue symptoms," "dengue treatment," and "dengue." The selection of search terms was informed by prior literature. The data were extracted at weekly granularity for the period January to December 2024, selecting Pakistan as the region of interest. Google Trends provides

Relative Search Volume (RSV) values from 0 to 100, with 100 representing the week that had the highest search frequency for the selected term.

Both datasets are publicly available and do not contain any personally identifiable information. Therefore, no ethical approval was needed.

Data Processing and Normalization
The IDSR case count data and Google Trends data were entered into Microsoft Excel and then imported into R version 4.3.3 for analysis.

Since Google Trends scores are scaled between 0 and 100 while case counts vary in absolute magnitude, the case count data were normalized to make both series comparable.
For each weekly value xi of dengue case counts, the normalized value y_i was computed as:

$$y_i = \frac{x_i}{x_{max}} \times 100$$

Where x_{max} is the maximum weekly case count in 2024.
This rescaling ensured that both datasets shared a common 0–100 scale.

Statistical Analysis
Correlational Analysis
All analyses were performed in R version 4.3.3 using the stats, zoo, and vars packages.
To assess the strength and direction of the relationship between Google search interest and dengue case count data, Pearson's product–moment correlation coefficient (r) was calculated separately for each search term. Two-tailed significance tests were conducted, and 95% confidence intervals (CI) were obtained.

Time-Series and Causality Analysis
Stationarity testing
Before causality testing, the Augmented Dickey–Fuller (ADF) test was applied to each time series to check for stationarity. Weekly Google Trends data for the search terms "dengue fever," "dengue symptoms," "dengue treatment," and "dengue" were analyzed alongside normalized dengue case count data from the IDSR system. A series was considered non-stationary when the ADF p-value > 0.05. Non-stationary series were differenced until stationarity was achieved.

The maximum differencing order dmax required across all series was used in the Toda–Yamamoto procedure.

Model specification and lag selection
A Vector Autoregression (VAR) model was constructed for each pairing of Google Trends data and normalized dengue case count data. The optimal lag length (k) was identified using multiple information criteria, including the Akaike Information Criterion (AIC), Hannan–Quinn (HQ), and Final Prediction Error (FPE) statistics. For the 2024 dataset, both AIC and HQ consistently indicated an optimal lag length of two weeks. This lag order was applied uniformly across all models to ensure consistency, comparability, and overall model stability. An augmented VAR with p=k+ dmax lags was used for the Toda–Yamamoto causality test.

Toda–Yamamoto causality test
The Toda–Yamamoto approach was employed to examine whether variations in Google search activity Granger-cause changes in dengue case counts. For each model, the null hypothesis stated that the Google Trends relative search volume (RSV) for a given term does not Granger-cause a rise in dengue case counts. Causality was evaluated using the modified Wald (MWALD) test applied to the coefficients of the augmented vector autoregressive (VAR) model. Statistical significance was defined at $p < 0.05$. The analysis generated F-statistics, degrees of freedom (df, df), and corresponding p-values for each model. In addition, instantaneous causality between variables was assessed using the same procedure to determine whether simultaneous relationships existed between Google search activity and dengue case count data.

Visualization
Weekly trends of each search term and dengue case count data were plotted, showing both raw and smoothed (3-week moving average) lines to illustrate synchronous or lagged patterns. Additionally, separate scatter plots were produced for each search term to depict linear correlations, with regression lines shown as dashed overlays.

Distinct color palettes were used to differentiate each search term: red ("dengue fever"), blue ("dengue symptoms"), green ("dengue treatment"), and purple ("dengue").

Results
Correlation between Google Trends data and dengue case counts
A strong positive correlation was observed between weekly Google search interest for dengue-related terms and normalized dengue case count data in Pakistan during 2024. As shown in Table 1, the Pearson correlation coefficients ranged from 0.756 to 0.832 ($p < 0.001$ for all). The highest correlation was seen for the general search term "dengue" ($r = 0.832$; 95% CI = 0.724–0.901), followed by "dengue symptoms" ($r = 0.786$; 95% CI = 0.653–0.872), "dengue treatment" ($r = 0.775$; 95% CI = 0.637–0.865), and "dengue fever" ($r = 0.756$; 95% CI = 0.608–0.853).

These results indicate that as public search activity for dengue-

related terms increased online, reported dengue case counts also rose. Line plots with 3-week moving averages (Figures 1–4) and scatter plots comparing Google Trends and IDSR data (Figure 5) both demonstrated clear synchronous patterns and positive linear relationships between online search interest and outbreak counts, reinforcing the observed correlations.

Causality analysis using the Toda–Yamamoto test
Before applying the Toda–Yamamoto causality test, the stationarity of each time series was examined using the Augmented Dickey–Fuller (ADF) test. The ADF test results indicated that all series were non-stationary in level form, with $p > 0.05$ ("dengue fever": Dickey–Fuller = −1.93; lag order = 3; $p = 0.60$; "dengue symptoms": Dickey–Fuller = −1.52; lag order = 3; $p = 0.77$; "dengue treatment": Dickey–Fuller = −1.43; lag order = 3; $p = 0.80$; "dengue": Dickey–Fuller = −2.75; lag order = 3; $p = 0.27$; IDSR: Dickey–Fuller = −1.57; lag order = 3; $p = 0.75$). After first differencing, all series remained non-stationary ($p > 0.05$ for all variables). However, after second-order differencing, the data became stationary ("dengue fever": Dickey–Fuller = −5.47; lag order = 3; $p < 0.01$; "dengue symptoms": Dickey–Fuller = −5.03; lag order = 3; $p < 0.01$; "dengue treatment": Dickey–Fuller = −5.35; lag order = 3; $p < 0.01$; "dengue": Dickey–Fuller = −3.62; lag order = 3; $p = 0.04$; IDSR: Dickey–Fuller = −5.19; lag order = 3; $p < 0.01$). The maximum differencing order required to achieve stationarity across all variables was dmax = 2.

No cointegration was observed between the Google Trends and IDSR series, confirming the appropriateness of the Toda–Yamamoto causality framework without requiring cointegration adjustments. A Vector Autoregressive (VAR) model was then constructed for each pairing of Google Trends data and IDSR case count data. The optimal lag length (k) was determined using multiple information criteria, including the AIC, HQ, and FPE statistics, all of which consistently indicated an optimal lag of two weeks. Accordingly, an augmented VAR model with p=k+ dmax = 4 was applied for each term.

Causality was assessed using the modified Wald (MWALD) test on the augmented VAR coefficients, with significance defined at $p < 0.05$. As shown in Table 2, significant Granger causality was observed for the search terms "dengue symptoms" ($F = 4.18$; $p = 0.004$), "dengue treatment" ($F = 5.10$; $p = 0.001$), and "dengue" ($F = 4.24$; $p = 0.004$), indicating that increases in online search activity preceded corresponding rises in reported dengue case counts. The observed time lag between Google search activity and reported dengue cases correlates well with the incubation periods of dengue virus transmission. The association for "dengue fever" was borderline ($F = 2.41$; $p = 0.056$). Instantaneous causality tests were non-significant across all models ($p > 0.05$), suggesting that relationships between search activity and case count data occurred with a time lag rather than simultaneously.

Discussion
This study looked at how Google Trends data could support Pakistan's national dengue surveillance system by tracking public search activity related to dengue over time. A strong positive relationship was found between weekly Google search interest and normalized case count data from the IDSR system, ($r = 0.76–0.83$, $p < 0.001$), with the term "dengue" demonstrating the strongest association. These findings indicate that as dengue transmission rises, people start searching online for information about illnesses, reflecting real-world disease trends in near real time.

The Toda-Yamamoto causality analysis showed that search activity can precede reported rises in case counts. Significant Granger causality appeared for "dengue symptoms" ($p = 0.004$), "dengue treatment" ($p = 0.001$), and "dengue" ($p = 0.004$), while "dengue fever" had a borderline association ($p = 0.056$). These findings suggest that rises in Google search interest usually happen weeks before officially reported rises in dengue cases. This supports the idea that digital search data could serve as an early warning tool for dengue surveillance in Pakistan.
Our findings align with research from other regions where dengue is common. In India, Singh et al. (2025) found a strong link between dengue-related Google searches and IDSP case counts ($r \approx 0.83$). Searches began to rise weeks before reported peaks [10,12]. Similarly, Verma et al. (2018) noted a lag of -2 to -3 weeks between search trends and surveillance data [9]. Other studies, including those by Sylvestre et al. and Ho et al., have shown similar correlations and time-lag patterns, usually ranging from 1 to 3 weeks [14,15]. These consistent time-lags suggest that public concern and the desire for information grow early in the outbreak—often during the phase when transmission starts to speed up but case reporting systems have not yet captured the increase. This closely matches our observed lag of 2 weeks and supports the idea that online search patterns can serve as real-time early indicators of transmission.

Comparable associations and predictive effects have also been reported in Indonesia (Husnayain et al., 2019), Brazil (Romero-Alvarez et al., 2020), and Mexico (Gluskin et al., 2014). This is especially true in urban areas where internet use is high and online health-seeking behavior is more common [11,16,17]. These patterns highlight an important contextual factor: the predictive value of Google Trends seems strongest in regions with high population density, active digital engagement, and frequent seasonal dengue outbreaks. These conditions are similar to major Pakistani cities like Karachi, Lahore, and Rawalpindi. Overall, these studies show a broader global trend: digital search activity not only reflects official rises in

case count data but often comes before it. This supports its role as a useful and scalable early-warning signal in national surveillance systems.

A major strength of this study is its demonstration of how digital epidemiology can significantly improve Pakistan's existing dengue surveillance system. Unlike traditional case reporting, which often faces delays due to laboratory confirmation, clinical triage capacity, and administrative processing, Google Trends data are publicly available, free, and updated in real time. This allows for earlier situational awareness. Importantly, our analysis used national IDSR surveillance data, ensuring that the correlations we observed reflect actual public health reporting rather than indirect indicators. The use of multiple search terms and both correlation and causality approaches, including Toda-Yamamoto testing, further strengthens our findings. This reduces the chance that the observed associations resulted from coincidence, short-term media spikes, or seasonal awareness alone. This method does not need extra laboratory or field resources, making it scalable and practical for routine integration into outbreak monitoring workflows. In places like Pakistan, where reporting delays, a substandard surveillance system and uneven access to diagnostics are ongoing challenges, this approach offers a useful, low-cost way to achieve earlier detection, quicker risk communication, entomological surveillance, trigger vector control, diagnostic preparedness and a more proactive response to outbreaks [8,9,18,19].

A key practical implication of this approach is its potential role in laboratory diagnostic preparedness. Rising Google search activity for terms such as "dengue symptoms" or "dengue treatment" may alert diagnostic laboratories to an impending increase in testing demand before official case reports peak. This could support early procurement of NS1 antigen kits, IgM/IgG ELISA reagents, RT-PCR supplies, sample collection materials, and blood count testing capacity. It may also help laboratories plan staffing, extend testing hours, prioritize turnaround time, and coordinate reporting pathways with public health authorities. Therefore, Google Trends should not be viewed only as an epidemiological signal, but as a low-cost trigger for laboratory readiness during dengue season.

At the same time, a few limitations should be considered. The observed approximate two-week lag between search activity and reported IDSR data should be interpreted cautiously. While it may partly reflect the clinical course of dengue and the interval between symptom onset and healthcare-seeking behavior, it is also likely influenced by delays in healthcare presentation, laboratory confirmation, and administrative reporting within Pakistan's passive surveillance system. Therefore, the observed lag probably represents a combination of biological and reporting factors rather than the incubation period alone.

Furthermore, Google Trends shows patterns in public search interest, which may be affected by media coverage or awareness campaigns, along with actual infection rates [20]. Additionally, internet access and digital skills are often more concentrated in urban areas. This means search activity might not fully show trends in rural or remote regions. Lastly, this study covered a single year period. While the 2024 weekly data captured seasonal trends, the relatively small sample size may reduce the stability of the VAR lag selection and Toda–Yamamoto causality tests. Consequently, these forecasting results should be treated as preliminary evidence of temporal associations rather than conclusive predictive outcomes. Looking at multi-year datasets covering several dengue seasons, and including more data sources like climate factors, vector numbers, or reports from provinces, would help further validate and improve the method. Overall, these points underline areas for future work while still supporting the usefulness of Google Trends as a helpful tool in dengue surveillance.

Conclusion

This study shows that Google Trends can be a useful and low-cost addition to Pakistan's dengue surveillance system. The strong link and clear time-lag relationship between search activity and reported case counts indicate that people's online behavior reflects real changes in disease spread and can even help predict them. By watching for increases in dengue-related searches, health authorities could receive early warning signals before hospital cases rise. While Google Trends shouldn't replace official reporting, combining it with existing systems like the IDSR and incorporating Google Trends with microbiological data such as RT-PCR positivity rates, NS1 antigen trends, and circulating serotypes could create a multilayered early warning system which could make outbreak monitoring faster and more responsive. Future studies should look at data from multiple years, include regional comparisons, and connect digital data with weather and mosquito surveillance to build better early warning models.

Table 1: Pearson correlation between weekly Google search interest for dengue-related terms and normalized IDSR dengue outbreak data in Pakistan (2024).

Search Term	Pearson's *r*	95% Confidence Interval	*p*-value	Strength of Correlation
Dengue Fever	0.756	0.608–0.853	< 0.001	Strong positive
Dengue Symptoms	0.786	0.653–0.872	< 0.001	Strong positive
Dengue Treatment	0.775	0.637–0.865	< 0.001	Strong positive
Dengue	0.832	0.724–0.901	< 0.001	Strong positive

Table 2: Toda–Yamamoto causality test results for weekly Google search interest and normalized dengue outbreak data in Pakistan (2024).

Search Term	F-statistic	df_1	df_2	p-value	Interpretation
Dengue Fever	2.41	4	78	0.056	Borderline causality (not statistically significant)
Dengue Symptoms	4.18	4	78	0.004	Significant causality (searches predict outbreaks)
Dengue Treatment	5.10	4	78	0.001	Significant causality (searches predict outbreaks)
Dengue	4.24	4	78	0.004	Significant causality (searches predict outbreaks)

Figure 1: Weekly Google Trends (GT) search interest for "dengue fever" and normalized IDSR dengue case count data in Pakistan (2024). The figure presents both raw values and 3-week moving averages for comparative trend analysis.

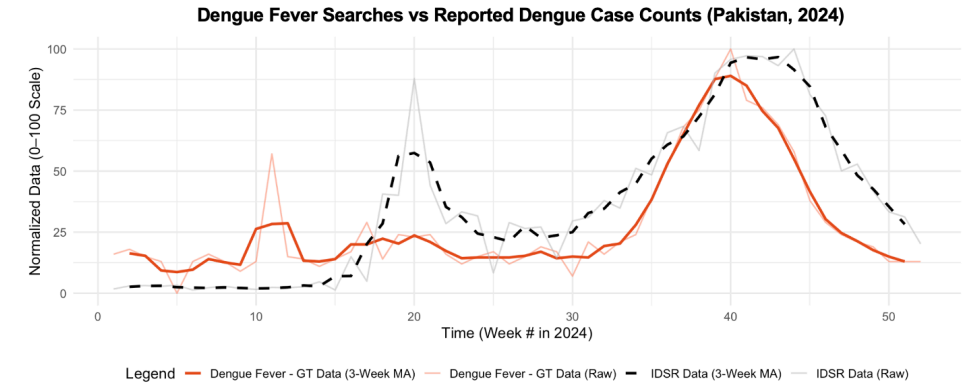


Figure 2: Weekly Google Trends (GT) search interest for "dengue symptoms" and normalized IDSR dengue case count data in Pakistan (2024). The figure presents both raw values and 3-week moving averages for comparative trend analysis.

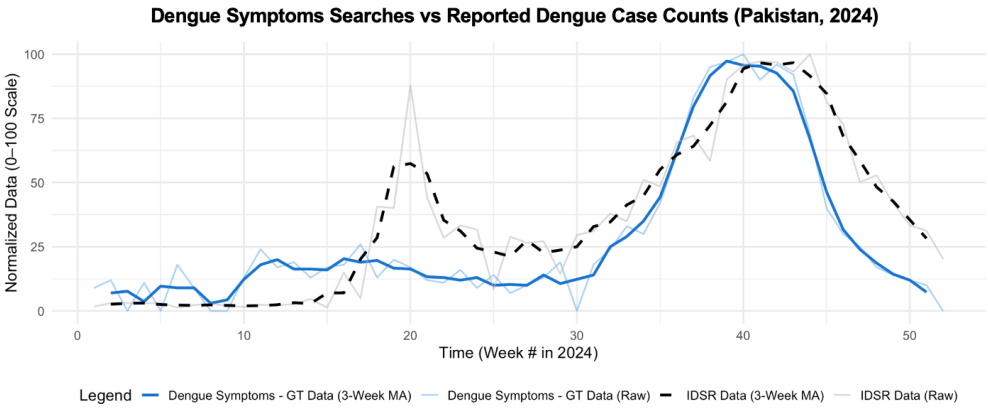


Figure 3: Weekly Google Trends (GT) search interest for "dengue treatment" and normalized IDSR dengue case count data in Pakistan (2024). The figure presents both raw values and 3-week moving averages for comparative trend analysis.

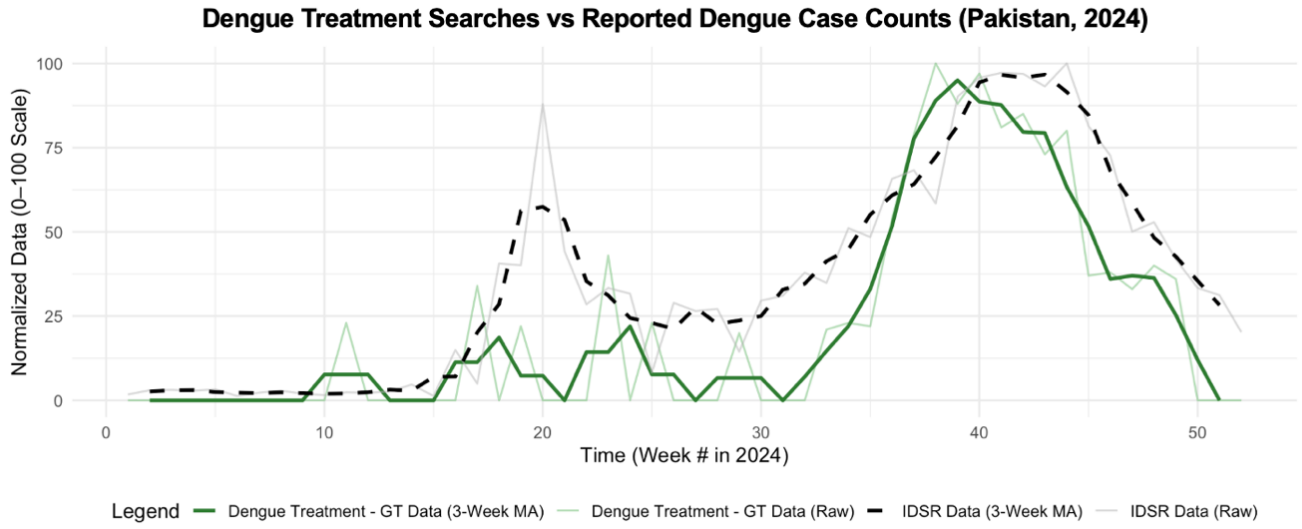


Figure 4: Weekly Google Trends (GT) search interest for "dengue" and normalized IDSR dengue case count data in Pakistan (2024). The figure presents both raw values and 3-week moving averages for comparative trend analysis.

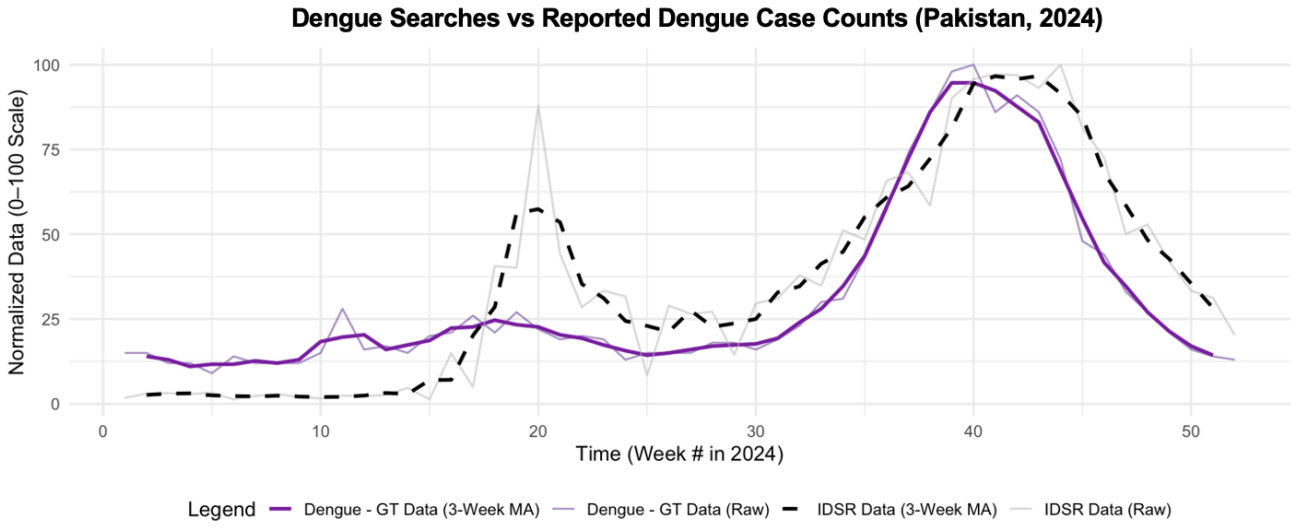
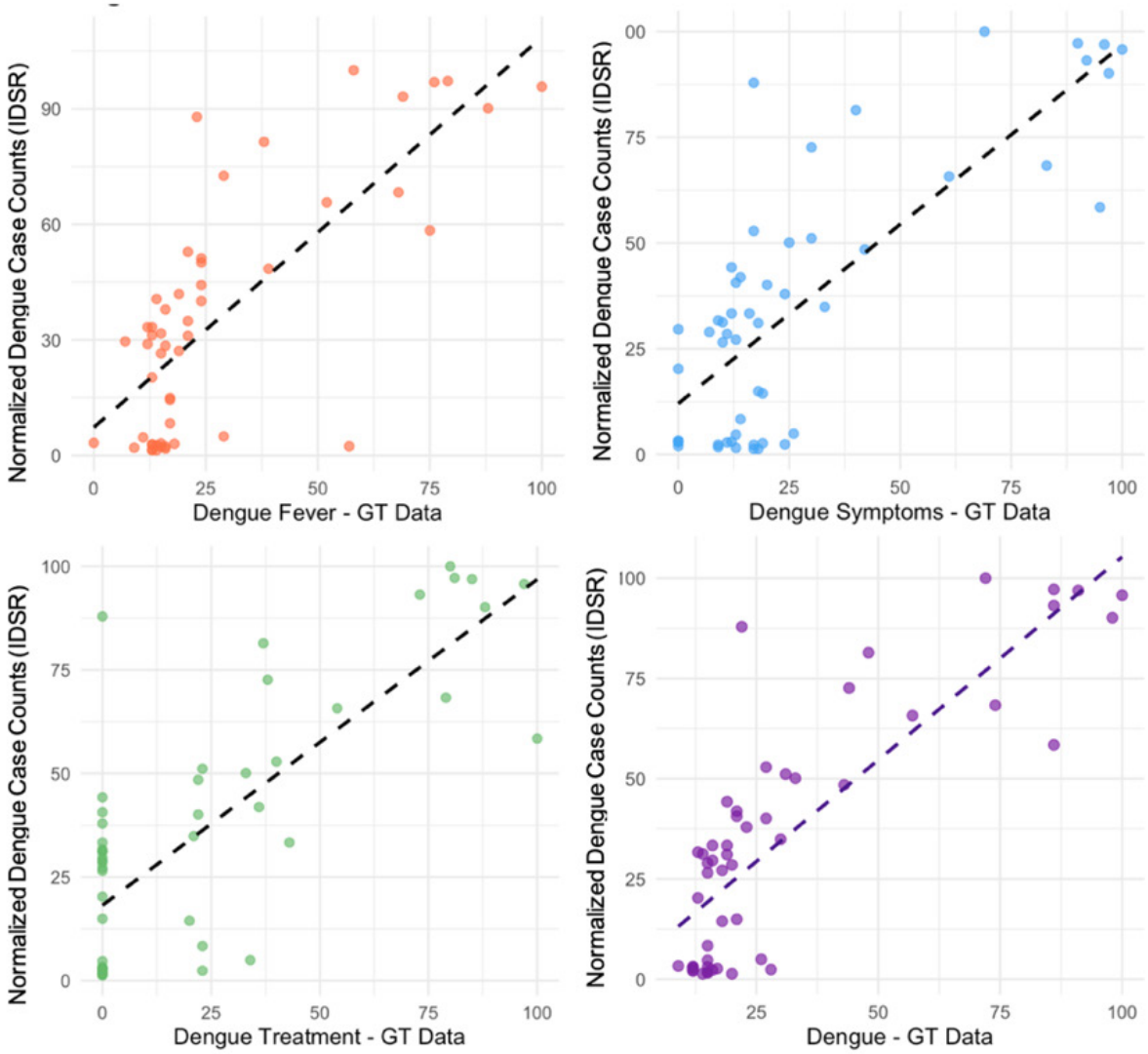


Figure 5: Scatter plots showing the relationship between Google Trends (GT) search interest and normalized dengue case count data (IDSR) in Pakistan during 2024. The top left, top right, bottom left, and bottom right panels represent plots for search terms: "Dengue Fever", "Dengue Symptoms", "Dengue Treatment", and "Dengue", respectively. The dashed line in each plot represents the linear regression line, illustrating the strong positive association between Google search activity and reported dengue case counts.



Conflicts of Interests
All authors declare no conflicts of interest.

Ethical Approval
Not required as publicly available data.

Credit Author Statements
Alizeh Sonia Fatimi did the data search, data analysis, and drafting of the manuscript. Tazeen Fatima, Asghar Nasir, Muhammad Shariq Shaikh, and Imran Siddiqui contributed to manuscript writing, graphs and critical revisions. Sibtain Ahmed conceived the primary idea for the work, supervised the overall project.

Funding
None.

Data Availability
Publicly available data was analysed.

References

1. Dengue worldwide overview [Internet]. Available from: https://www.ecdc.europa.eu/en/dengue-monthly. Accessed: 03/11/2025.
2. Guzman MG, Harris E. Dengue. Lancet. 2015;385(9966):453–465.
3. Khan J, Khan I, Ghaffar A, Khalid B. Epidemiological trends and risk factors associated with dengue disease in Pakistan (1980–2014): a systematic literature search and analysis. BMC Public Health. 2018;18(1):745.
4. National Institute of Health, Pakistan. Advisory for the prevention and control of dengue fever. Islamabad: Center for Disease Control, Ministry of National Health Services, Regulations & Coordination; 2025.
5. Rafique I, Nadeem Saqib MA, Munir MA, Qureshi H, Siddiqui S, Habibullah S, et al. Economic burden of dengue in four major cities of Pakistan during 2011. J Pak Med Assoc. 2015;65(3):256–259.
6. Hussain M, Majeed A, Imran I, Rasool MF, Hussain I, Ghaffar A, et al. Dengue control in Pakistan: prior planning is better than controlling too late. BMJ. 2019;367:l6912.
7. Haq ZU, Fazid S, Hussain B, Khan MF, Betanni A, Behrawar B, et al. Assessment of the impact of integrated disease surveillance and response system on surveillance management at healthcare facilities in Pakistan. East Mediterr Health J. 2024;30(2):109–115.
8. Shaikh TG, Waseem S, Ahmed SH, Swed S, Hasan MM. Infectious disease surveillance system in Pakistan: challenges and way forward. Trop Med Health. 2022;50(1):46.
9. Verma M, Kishore K, Kumar M, Sondh AR, Aggarwal G, Kathirvel S. Google search trends predicting disease outbreaks: an analysis from India. Healthc Inform Res. 2018;24(4):300–308.
10. Thamaraikannan R, Munisamy V. Using Google Trends for dengue surveillance and epidemiological research. Int J Community Med Public Health. 2022;9(10):3917.
11. Husnayain A, Fuad A, Lazuardi L. Correlation between Google Trends on dengue fever and national surveillance report in Indonesia. Glob Health Action. 2019;12(1):1552652.
12. Singh G, Jha A, Aadil MK. Analyzing the relationship between Google Trends data and dengue outbreaks: a causality and correlation study. Discover Public Health. 2025;22(1):320.
13. Individuals using the Internet (% of population) – Pakistan [Internet]. Available from: https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=PK. Accessed: 03/11/2025.
14. Ho HT, Carvajal TM, Bautista JR, Capistrano JDR, Viacrusis KM, Hernandez LFT, et al. Using Google Trends to examine the spatio-temporal incidence and behavioral patterns of dengue disease: a case study in Metropolitan Manila, Philippines. Trop Med Infect Dis. 2018;3(4):118.
15. Sylvestre E, Cécilia-Joseph E, Bouzillé G, Najioullah F, Etienne M, Malouines F, et al. The role of heterogeneous real-world data for dengue surveillance in Martinique: observational retrospective study. JMIR Public Health Surveill. 2022;8(12):e37122.
16. Gluskin RT, Johansson MA, Santillana M, Brownstein JS. Evaluation of Internet-based dengue query data: Google Dengue Trends. PLoS Negl Trop Dis. 2014;8(2):e2713.
17. Romero-Alvarez D, Parikh N, Osthus D, Martinez K, Generous N, del Valle S, et al. Google Health Trends performance reflecting dengue incidence for the Brazilian states. BMC Infect Dis. 2020;20(1):252.
18. Wang D, Guerra A, Wittke F, Lang JC, Bakker K, Lee AW, et al. Real-time monitoring of infectious disease outbreaks with a combination of Google Trends search results and the moving epidemic method: a respiratory syncytial virus case study. Trop Med Infect Dis. 2023;8(2):75.
19. Satpathy P, Kumar S, Prasad P. Suitability of Google Trends for digital surveillance during ongoing COVID-19 epidemic: a case study from India. Disaster Med Public Health Prep. 2021;17:e28.
20. Chan EH, Sahai V, Conrad C, Brownstein JS. Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. PLoS Negl Trop Dis. 2011;5(5):e1206.

Research Article

Population-Specific Reference Intervals for Whole-Blood GSH/GSSG Ratio in Healthy Emirati Adults According to CLSI and ISO 15189 Guidelines

Sushma Ravindra Umrao[1], Sampathkumar Elangovan[1*], Muhammad Raza Qureshi[1], Muhammad Masoud Alam[1], Farhan Ashraf[1], Shafaath Husain Syed Mohammed[1], Mohamed Sayed Tharammal[1], Faiz Ul Amin[1], Mohammed Munir Khan[1]

[1]*Life Core Private Clinic, Abu Dhabi, United Arab Emirates*

Article Info

[*]***Corresponding Author:***
Sampathkumar Elangovan
Life Core Private Clinic,
Royal Group Headquarters (7th Floor), Al Salam Street,
PO Box 63811, Abu Dhabi, United Arab Emirates
E-mail: Sampath.kumar@royalgroupuae.com

Keywords

Glutathione, GSSG, Reference interval, UHPLC, Oxidative stress, Emirati population, CLSI C28-A3, ISO 15189

Abstract

Background: The glutathione redox system, expressed as the ratio of reduced glutathione (GSH) to oxidized glutathione (GSSG), is a key indicator of cellular oxidative status. Manufacturer-provided reference intervals are typically derived from non-local populations and may not be applicable across different ethnic and geographic groups.

Objective: To characterize the distribution of whole-blood GSH/GSSG ratio in healthy Emirati adults using ultra-high-performance liquid chromatography (UHPLC) and to establish population-specific reference intervals in accordance with CLSI C28-A3 and ISO 15189 guidelines.

Methods: In this cross-sectional study, whole-blood samples from 120 apparently healthy Emirati adults were analyzed using a commercial UHPLC assay. The GSH/GSSG ratio was calculated for each participant. Due to non-normal distribution, reference intervals were derived using a nonparametric approach. Both the conventional 95% reference interval (2.5th–97.5th percentiles) and the central 90% interval (5th–95th percentiles) were calculated. Bootstrap resampling (20,000 iterations) was used to estimate 95% confidence intervals for the central 90% limits.

Results: The GSH/GSSG ratio demonstrated a right-skewed distribution with a median (IQR) of 4.76 (3.58–6.77). Most participants (89.2%) had values below the manufacturer-reported reference interval (10–15), with only 5.0% within this range. The nonparametric 95% reference interval was 1.66–28.40. The central 90% interval was 2.53–17.20, with bootstrap 95% confidence intervals of 2.21–2.97 (lower) and 14.92–19.73 (upper).

Conclusion: Manufacturer-provided reference intervals do not reflect the distribution observed in healthy Emirati adults. Population-specific reference intervals are therefore necessary. Reporting both 95% and 90% intervals, supported by bootstrap estimates, enhances clinical interpretation for both diagnostic and screening applications.

Introduction

Oxidative stress, defined as an imbalance between the production of reactive oxygen species and antioxidant defense capacity, plays a central role in the pathophysiology of numerous chronic and age-related diseases, including cardiovascular disease, diabetes, neurodegeneration, and cancer [1,2]. Maintenance of cellular redox homeostasis is therefore essential for normal physiological function, and reliable biomarkers of redox status are increasingly important in both clinical and research settings.

Among endogenous antioxidant systems, the glutathione redox couple, consisting of reduced glutathione (GSH) and oxidized glutathione (GSSG), represents a primary regulator of intracellular redox balance [3]. The GSH/GSSG ratio is widely recognized as a sensitive and integrative indicator of oxidative stress, reflecting both antioxidant capacity and oxidative burden [4,5]. Unlike absolute concentrations of GSH or GSSG alone, the ratio provides insight into redox dynamics and has been associated with metabolic, inflammatory, and degenerative conditions [6,7].

Despite its recognized biological relevance, clinical interpretation of the GSH/GSSG ratio remains challenging, as many assays rely on manufacturer-derived or externally established reference intervals that may not be fully representative of the populations in which they are applied [8]. Increasing evidence suggests that redox biomarkers, including glutathione-related parameters, exhibit significant inter-population variability, influenced by ethnicity, diet, lifestyle, environmental exposures, and genetic background [9,10]. Application of non-local reference intervals may therefore lead to systematic misclassification of healthy individuals.

International laboratory standards explicitly recognize this limitation. The Clinical and Laboratory Standards Institute (CLSI) C28-A3 guideline emphasizes the need for population-specific reference interval verification or establishment, particularly when demographic or ethnic differences are expected [11]. Similarly, ISO 15189 requires clinical laboratories to ensure that reference intervals are appropriate for the population served and to justify their clinical applicability [12]. However, population-specific reference intervals for redox biomarkers, including the GSH/GSSG ratio, remain scarce, particularly in Middle Eastern populations.

To date, no published study has established whole-blood GSH/GSSG reference intervals for healthy Emirati adults using modern chromatographic techniques. Given the unique genetic, dietary, and environmental characteristics of the Emirati population, extrapolation from reference intervals derived in Western or East Asian cohorts may be inappropriate. Furthermore, advances in ultra-high-performance liquid chromatography (UHPLC) offer improved analytical specificity and sensitivity for glutathione measurement compared with older spectrophotometric or enzymatic methods [13].

The present study was therefore designed to address this knowledge gap by characterizing the distribution of whole-blood GSH/GSSG ratios in a cohort of apparently healthy Emirati adults and establishing population-specific reference intervals in accordance with CLSI C28-A3 and ISO 15189 guidelines. In addition to the conventional 95% reference interval, a central 90% interval was derived to support sensitivity-oriented screening and early redox imbalance detection, providing laboratories with greater interpretive flexibility.

Materials and Methods

Study design and participants

This cross-sectional study included 120 apparently healthy Emirati adults. Participants were recruited from the local community and reported no known acute or chronic illness at the time of sampling. As the study involved exclusively anonymized retrospective data, formal Institutional Review Board approval and ethical committee registration were not required in accordance with institutional policy and applicable regulations, and the requirement for informed consent was waived. The study was conducted in accordance with institutional data protection standards and the principles of the Declaration of Helsinki.

Sample collection and UHPLC analysis

Whole-blood samples were collected by venipuncture into anticoagulant-containing tubes and processed according to the manufacturer's instructions to preserve glutathione redox status. Reduced glutathione (GSH) and oxidized glutathione (GSSG) concentrations were measured using the Chromsystems "Glutathione in Whole Blood" HPLC reagent kit (Order No. 66000; Chromsystems Instruments & Chemicals GmbH, Gräfelfing, Germany), according to the manufacturer's instructions provided in the assay manual (Instruction Manual IM 66000 Glutathione EN 06/2016 R3). For pre-analytical handling, fresh whole-blood samples were processed immediately after collection. Stored frozen samples were thawed immediately prior to analysis and processed without delay together with calibrators and quality controls. Sample preparation was performed under chilled and light-protected conditions to minimize ex vivo oxidation of glutathione and alterations in the GSH/GSSG ratio. Repeated freeze–thaw cycles were avoided. The GSH/GSSG ratio was calculated for each participant.

Statistical analysis

Statistical analyses were performed using SPSS Statistics (IBM Corp., Armonk, NY, USA). Descriptive statistics were reported as mean (SD) and median (IQR). Distributional characteristics were assessed using histograms, boxplots, empirical cumulative distribution functions, and quantile–quantile plots. Due to non-normality, reference intervals were derived using a nonparametric percentile-based approach in accordance with CLSI C28-A3 guidelines.

Both a 95% reference interval (2.5th–97.5th percentiles)

and a central 90% interval (5th–95th percentiles) were calculated to support diagnostic and screening-oriented clinical interpretation, consistent with ISO 15189 laboratory practice. Observed values were compared with the manufacturer-provided reference interval (10–15). Bootstrap resampling (20,000 iterations) was used to estimate 95% confidence intervals for the central 90% reference limits. Outliers were identified using Tukey's criterion (Q3 + 1.5×IQR) and retained in the primary analysis.

Results

1. Study population and data completeness

Whole-blood GSH/GSSG ratio measurements were available for all 120 apparently healthy Emirati adults, and no missing or invalid data points were identified. All samples were included in the primary analysis.

2. Distributional characteristics of the GSH/GSSG ratio

The GSH/GSSG ratio demonstrated a right-skewed, non-normal distribution, with most values clustered at the lower end and a small proportion of high-value observations. The median (IQR) ratio was 4.76 (3.58–6.77), while the mean (SD) was 6.45 (5.51). Observed values ranged from 1.44 to 30.23, indicating substantial inter-individual variability.

Normality assessment using quantile–quantile plots confirmed deviation from Gaussian distribution on the raw scale, with partial improvement following logarithmic transformation (Figures 1 and 2).

Table 1 summarizes the descriptive statistics for the GSH/GSSG ratio in the study population.

Table 1: Distribution characteristics and descriptive statistical summary of whole-blood GSH/GSSG ratio values in apparently healthy Emirati adults (n = 120).

Parameter	Value
Mean (SD)	6.45 (5.51)
Median (IQR)	4.76 (3.58–6.77)
Minimum – Maximum	1.44 – 30.23
Skewness	2.89
Excess kurtosis	8.80

Figure 1: Normal Q-Q plot of raw GSH/GSSG ratio values. Quantile-quantile plot showing substantial deviation from normality, particularly in the upper tail.

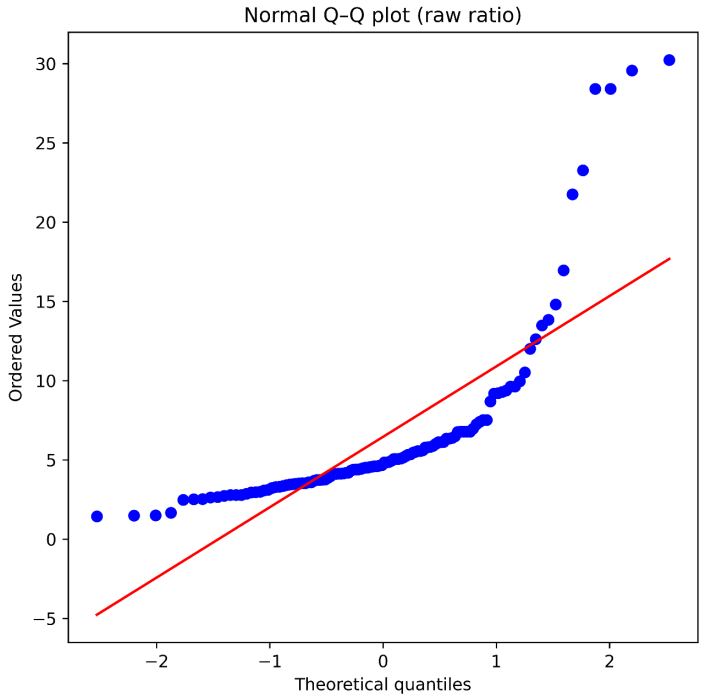
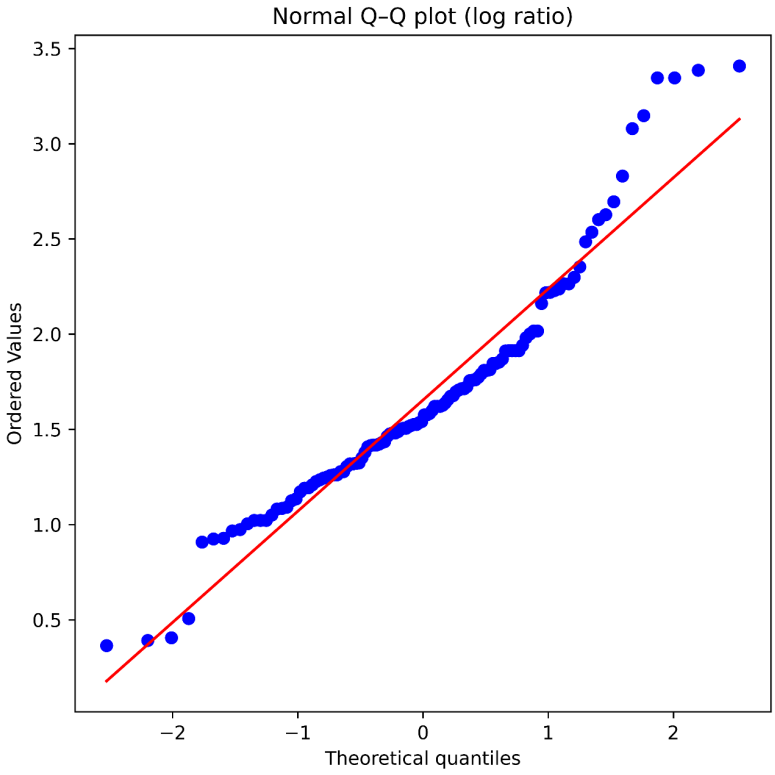
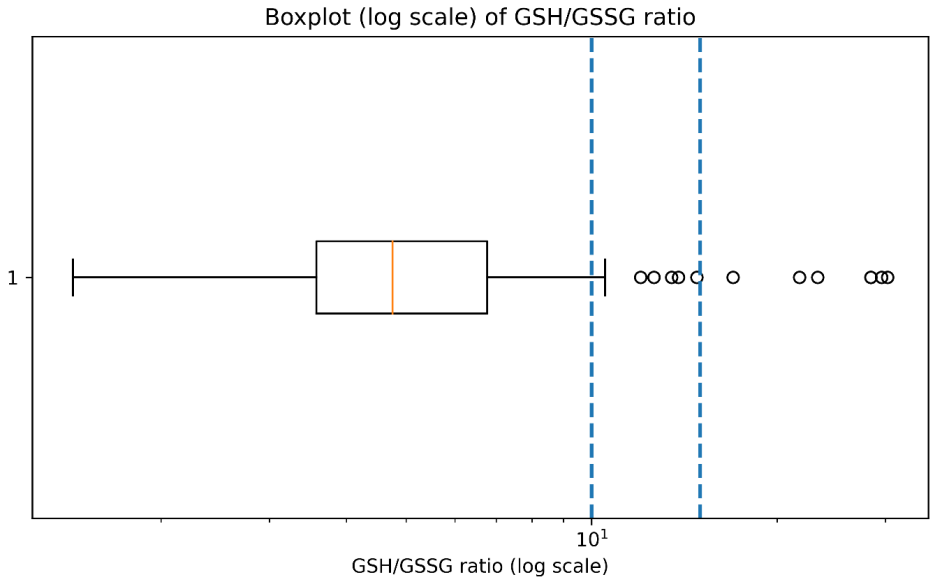


Figure 2: Normal Q-Q plot of log-transformed GSH/GSSG ratio values. Log transformation partially improves normality, but residual deviations persist, supporting the use of nonparametric methods.



The distribution and presence of high-value observations are further illustrated using a logarithmic-scale boxplot (Figure 3), which highlights right-skewness and inter-individual variability across the observed range.

Figure 3: Log-scale boxplot of whole-blood GSH/GSSG ratio in healthy Emirati adults. Horizontal boxplot displayed on a logarithmic scale illustrating the median, interquartile range, and high-value outliers of the GSH/GSSG ratio (n = 120). The log transformation highlights right-skewness and inter-individual variability in redox status.



3. Comparison with manufacturer-provided reference interval

The manufacturer-reported reference interval for the GSH/GSSG ratio (10–15) showed poor concordance with the observed distribution. A large majority of participants (89.2%) had values below the vendor's lower limit, while only 5.0% fell within the stated reference interval and 5.8% exceeded the upper limit.

This discrepancy is visually evident in the histogram and boxplot, where the vendor reference band lies above the central tendency of the population distribution (Figures 4 and 5). The categorical comparison with the manufacturer's reference interval is presented in Table 2.

Figure 4: Distribution of whole-blood GSH/GSSG ratio in healthy Emirati adults. Histogram with kernel density overlay showing the distribution of GSH/GSSG ratio values (n = 120). The shaded band represents the manufacturer-provided reference interval (10–15). Dashed vertical lines indicate the 2.5th and 97.5th percentiles derived from the study population.

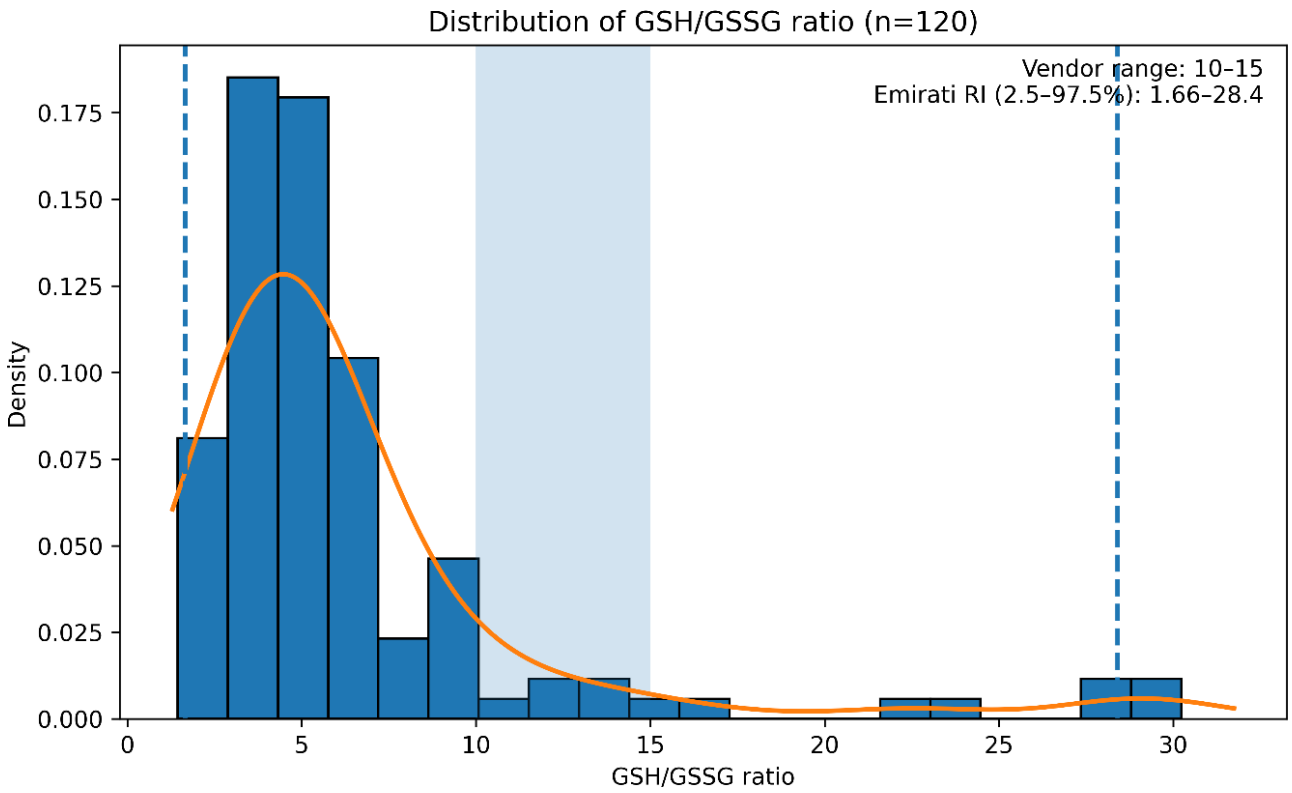


Figure 5: Boxplot of GSH/GSSG ratio relative to the manufacturer reference interval. Horizontal boxplot illustrating median, interquartile range, and outliers. The shaded region denotes the vendor reference interval (10–15), highlighting the predominance of values below this range.

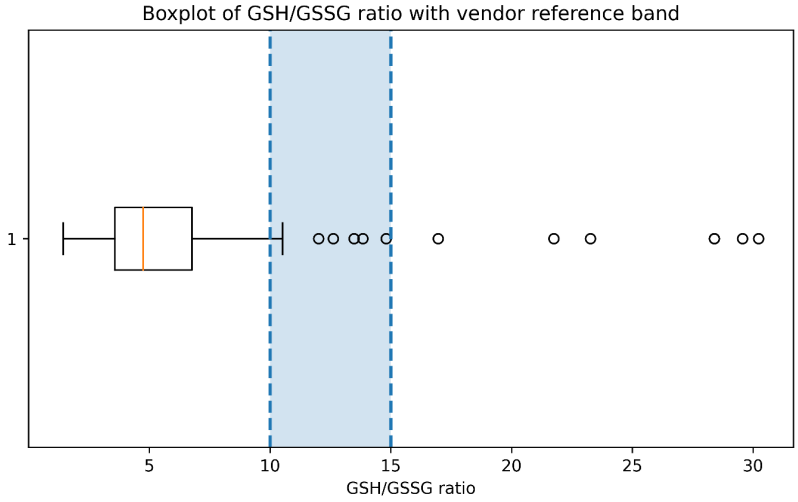
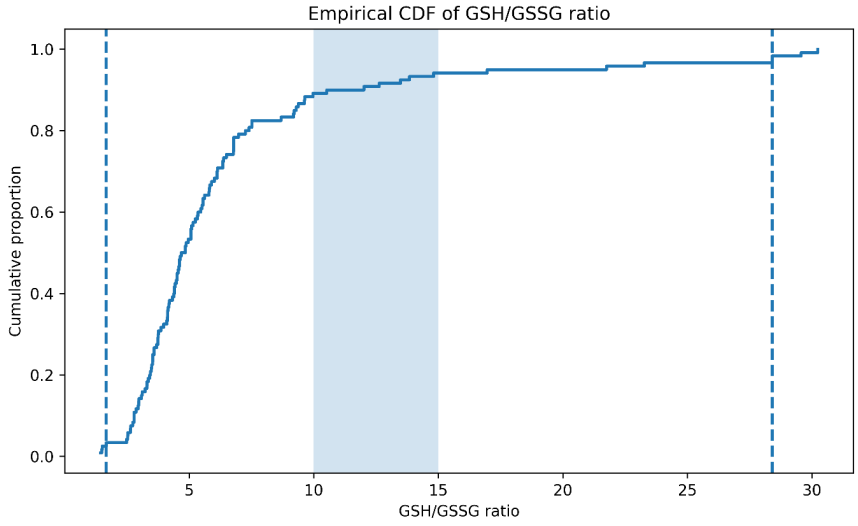


Table 2: Classification of whole-blood GSH/GSSG ratio values in healthy Emirati adults relative to the manufacturer-provided reference interval (10–15).

Classification relative to vendor range	Number (%)
Below 10	107 (89.2%)
Within 10–15	6 (5.0%)
Above 15	7 (5.8%)

The empirical cumulative distribution function (ECDF) further demonstrates this discordance by showing that approximately 90% of healthy individuals have GSH/GSSG ratio values below the manufacturer's lower cutoff of 10 (Figure 6).

Figure 6: Empirical cumulative distribution function (ECDF) of whole-blood GSH/GSSG ratio. The ECDF depicts the cumulative proportion of individuals across the observed range of GSH/GSSG ratio values. The curve demonstrates that the majority of healthy participants fall below the manufacturer-provided reference interval (10-15), supporting the derivation of population-specific reference intervals.

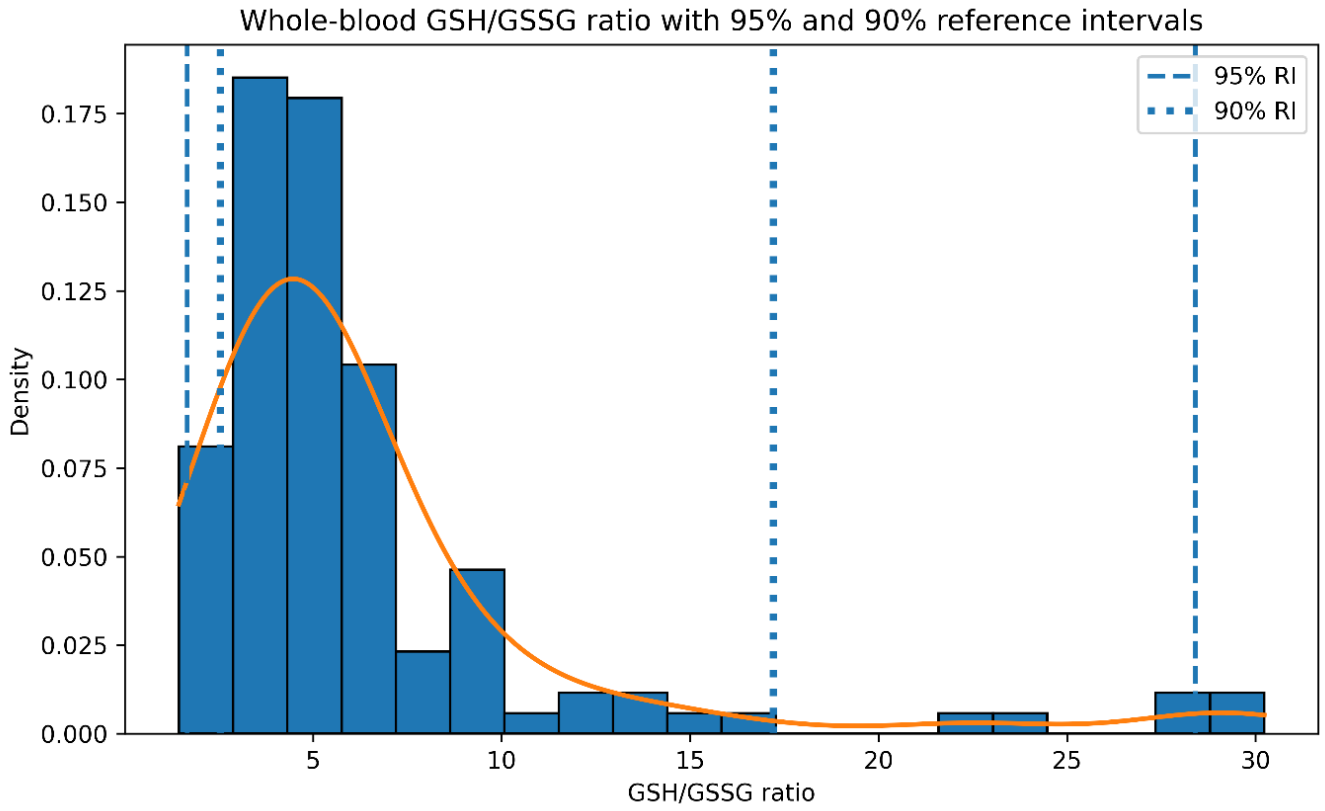


4. Population-specific reference interval estimation
Using a nonparametric percentile-based approach consistent with CLSI C28-A3 recommendations, a population-specific 95% reference interval (2.5th–97.5th percentiles) for the whole-blood GSH/GSSG ratio was derived as 1.66-28.40.
To support sensitivity-oriented screening and early redox imbalance detection, a central 90% reference interval (5th–95th percentiles) was also calculated, yielding limits of 2.53–17.20. The relationship between the observed distribution and both reference intervals is illustrated in the dual-interval distribution plot (Figure 7), where dashed lines represent the 95% reference interval and dotted lines indicate the central 90% interval. Both reference intervals are summarized in Table 3.

Table 3: Population-specific nonparametric diagnostic and screening reference intervals for whole-blood GSH/GSSG ratio in healthy Emirati adults derived according to CLSI C28-A3 guidelines.

Interval type	Percentiles	Lower limit	Upper limit	Intended clinical use
Reference interval (95%)	2.5th–97.5th	1.66	28.40	Diagnostic specificity
Central interval (90%)	5th–95th	2.53	17.20	Screening / early detection

Figure 7: Distribution of whole-blood GSH/GSSG ratio showing diagnostic and screening reference intervals. Histogram with kernel density overlay illustrating the distribution of GSH/GSSG ratio values in healthy Emirati adults (n = 120). Dashed vertical lines indicate the 95% reference interval (2.5th–97.5th percentiles), while dotted lines denote the central 90% interval (5th–95th percentiles). The dual-interval visualization highlights the distinction between diagnostic specificity and sensitivity-oriented screening thresholds.



5. Bootstrap confidence intervals for central 90% limits
Bootstrap resampling (20,000 iterations) was performed to assess the precision of the central 90% reference limits. The 95% confidence interval for the lower 5th percentile was 2.21–2.97, while the 95% confidence interval for the upper 95th percentile was 14.92–19.73, indicating acceptable robustness of the screening-oriented limits despite the skewed distribution. Bootstrap-derived confidence intervals are reported in Table 4 and contextualized visually in Figure 5.

Table 4: Bootstrap-derived 95% confidence intervals for the lower and upper limits of the central 90% whole-blood GSH/GSSG reference interval.

Reference limit	Point estimate	Bootstrap 95% CI
Lower limit (5th percentile)	2.53	2.21 – 2.97
Upper limit (95th percentile)	17.20	14.92 – 19.73

6. Outlier assessment and sensitivity analysis

Twelve participants were identified as high-value outliers based on Tukey's criterion (values > Q3 + 1.5×IQR). Exclusion of these individuals resulted in a narrower upper reference limit; however, the central tendency of the distribution and the observed discordance with the manufacturer-provided reference interval remained unchanged.

Accordingly, all participants were retained in the primary analysis, and both reference intervals were reported to reflect real-world biological variability in a healthy population.

Discussion

In this cross-sectional study of 120 apparently healthy Emirati adults, we established population-specific reference intervals for the whole-blood GSH/GSSG ratio using a validated UHPLC-based assay. The principal findings were: (i) the GSH/GSSG ratio exhibited a right-skewed, non-Gaussian distribution; (ii) the majority of participants had ratio values substantially lower than manufacturer-provided reference intervals; and (iii) the derived nonparametric 95% and central 90% reference intervals differed markedly from vendor-supplied limits derived from non-local populations.

The observed distribution of GSH/GSSG ratios in this cohort is consistent with prior reports indicating substantial biological variability in redox biomarkers among healthy individuals [9,10]. The predominance of lower ratio values, with a smaller subset of high-value observations, resulted in a right-skewed distribution, supporting the use of nonparametric methods for reference interval estimation as recommended by CLSI C28-A3.

A key finding of this study was the poor concordance between observed values and the manufacturer-provided reference interval. Nearly 90% of participants had GSH/GSSG ratios below the vendor's lower cutoff, indicating that application of the manufacturer's range would classify the majority of healthy Emirati adults as biochemically abnormal. Similar discrepancies have been reported in other populations when reference intervals derived from external cohorts are applied without local validation [8]. These findings underscore the limitations of universal reference intervals for redox biomarkers and reinforce the need for population-specific validation.

Comparison with previous population-based studies further highlights inter-population variability. Studies conducted in European and East Asian cohorts have reported higher median or mean GSH/GSSG ratios in whole blood or erythrocytes, although direct comparison is complicated by differences in analytical platforms, sample matrices, and reporting conventions [13,6]. Ethnic background, dietary antioxidant intake, environmental exposures, and lifestyle factors such as physical activity and smoking prevalence may all contribute to these differences [10,1]. Additionally, analytical methodology plays a critical role; UHPLC-based techniques generally provide greater specificity than older enzymatic or colorimetric assays, which may overestimate GSH or underestimate GSSG [13].

Several biological and methodological factors may explain the lower GSH/GSSG ratios observed in the present cohort. Diets characteristic of the region, differences in micronutrient intake, and environmental stressors such as heat exposure may influence baseline redox balance [2]. From an analytical perspective, whole-blood measurement captures both intracellular and extracellular glutathione pools, potentially yielding different distributions compared with erythrocyte-only or plasma-based assays [5].

An important strength of this study is the dual reporting of reference intervals. While the 95% reference interval remains the standard for diagnostic specificity, the central 90% interval may offer enhanced sensitivity for screening applications and early detection of redox imbalance. This approach is consistent with ISO 15189 guidance, which emphasizes clinically meaningful interpretation over rigid adherence to a single cutoff. The use of bootstrap-based confidence intervals further strengthens the robustness and transparency of the derived limits. From a clinical laboratory perspective, adoption of these population-specific intervals may improve the interpretation of GSH/GSSG results in Emirati adults by reducing false classification of healthy individuals as having abnormal redox status when non-local manufacturer-provided reference intervals are used. These intervals may also support more appropriate reporting, clinical decision-making, and longitudinal monitoring of oxidative stress–related conditions.

For laboratories operating under ISO 15189 standards, locally derived or locally verified reference intervals may provide stronger evidence that reported interpretive limits are appropriate for the population served. However, laboratories adopting these intervals should ensure that similar analytical methods, sample type, and pre-analytical handling procedures are used before implementation.

To our knowledge, this study represents the first establishment

of population-specific whole-blood GSH/GSSG reference intervals in healthy Emirati adults using UHPLC. The findings have direct implications for clinical laboratories and researchers using redox biomarkers in Middle Eastern populations. Adoption of locally derived reference intervals may reduce misclassification, improve clinical interpretation, and support more accurate assessment of oxidative stress–related conditions.

Several limitations should be acknowledged. The study cohort, while sufficient for nonparametric reference interval estimation, was limited to apparently healthy adults and did not include stratification by age, sex, or lifestyle factors. Future studies in larger cohorts should explore demographic stratification and longitudinal variability. Nonetheless, the present findings provide a necessary foundation for laboratory implementation and further research.

In conclusion, the GSH/GSSG ratio in healthy Emirati adults differs substantially from manufacturer-provided reference intervals derived from non-local populations. Establishment of population-specific reference intervals, supported by robust analytical methods and international guidelines, is essential for accurate clinical interpretation of redox biomarkers.

Conclusion

In conclusion, this study established population-specific reference intervals for the whole-blood GSH/GSSG ratio in apparently healthy Emirati adults using a UHPLC-based analytical approach. The observed distribution of GSH/GSSG ratios differed substantially from manufacturer-provided reference intervals derived from non-local populations, with the majority of healthy individuals exhibiting lower ratio values. These findings demonstrate that unverified application of external reference intervals may result in systematic misclassification when interpreting redox biomarkers.

By deriving both the conventional 95% reference interval and a central 90% interval in accordance with CLSI C28-A3 and ISO 15189 recommendations, this study provides laboratories with flexible, clinically meaningful tools for diagnostic and screening-oriented interpretation of redox status. The inclusion of bootstrap-based confidence intervals further strengthens the robustness and transparency of the proposed limits.

As the first study to report whole-blood GSH/GSSG reference intervals in healthy Emirati adults, this work addresses an important knowledge gap and highlights the necessity of population-specific validation for redox biomarkers. Future studies in larger and more diverse cohorts are warranted to explore demographic stratification, longitudinal variability, and clinical applicability in disease settings.

Acknowledgements

None.

Research Ethical

This study was conducted as a retrospective analysis of fully deidentified data obtained from routine laboratory records. In accordance with institutional policy and applicable regulations, research involving exclusively anonymized data does not require formal Institutional Review Board (IRB) approval or ethical committee registration, and the requirement for informed consent was therefore waived.

The study was conducted in compliance with institutional data protection standards and in accordance with the principles of the Declaration of Helsinki.

Author contributions

The authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest

The authors do not have any conflict of interest.

Funding

This study did not receive any specific funding.

Data availability statement

The data generated and analyzed in the presented study are available from the corresponding author on request.

References

1. Sies H. Oxidative stress: a concept in redox biology and medicine. Redox Biol. 2015;4:180-183. DOI: 10.1016/j.redox.2015.01.002
2. Jones DP. Redox theory of aging. Redox Biol. 2015;5:71-79. doi: 10.1016/j.redox.2015.03.004.
3. Lu SC. Glutathione synthesis. Biochim Biophys Acta. 2013;1830:3143-3153. doi: 10.1016/j.bbagen.2012.09.008.
4. Jones DP. Redefining oxidative stress. Antioxid Redox Signal. 2006;8:1865-1879. DOI: 10.1089/ars.2006.8.1865
5. Pastore A, Piemonte F, Locatelli M, Lo Russo A, Gaeta LM, Tozzi G, et al. Determination of blood total, reduced, and oxidized glutathione in pediatric subjects. Clin Chem. 2001;47:1467-1469. doi.org/10.1093/clinchem/47.8.1467
6. Nuttall SL, Martin U, Sinclair AJ, Kendall MJ. Glutathione: in sickness and in health. Lancet. 1998;351:645-646. doi: 10.1016/s0140-6736(05)78428-2.
7. Giustarini D, Colombo G, Garavaglia ML, Portinaro NM, Reggiani F, Mignini F, et al. Assessment of glutathione/glutathione disulfide ratio and S-glutathionylated proteins in human samples. Free Radic Biol Med. 2017;112:360-375. DOI: 10.1016/j.freeradbiomed.2017.08.008
8. Ozarda Y, Higgins V, Adeli K. Verification of reference intervals in routine clinical laboratories: practical challenges and recommendations. Clin Chem Lab Med. 2019;57:30–37. DOI: 10.1515/cclm-2018-0059
9. Giustarini D, Tsikas D, Colombo G, Milzani A, Dalle-Donne I, Fanti P, et al. Pitfalls in the analysis of the physiological antioxidant glutathione and its disulfide in biological samples. J Chromatogr B Analyt Technol Biomed Life Sci. 2016;1019:21-28. doi: 10.1016/j.

jchromb.2016.02.015.
10. Margaritelis NV, Nastos GG, Vasileiadou O, Chatzinikolaou PN, Theodorou AA, Paschalis V, et al. Inter-individual variability in redox and performance responses after antioxidant supplementation. Acta Physiol (Oxf). 2023;238:e14017. doi: 10.1111/apha.14017
11. CLSI (Clinical and Laboratory Standards Institute). Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline - third edition (CLSI C28-A3). Wayne (PA): CLSI; 2008.
12. International Organization for Standardization. Medical laboratories - requirements for quality and competence (ISO 15189). Geneva: ISO; 2012.
13. Giustarini D, Dalle-Donne I, Milzani A, Fanti P, Rossi R. Analysis of GSH and GSSG after derivatization with N-ethylmaleimide. Nat Protoc. 2013;8:1660-1669. DOI: 10.1038/nprot.2013.095

Research Article

Impact of Total Protein and Triglycerides on the comparability of Sodium and Potassium Measurements by Direct and Indirect Ion-Selective Electrodes : An Experimental Analytical Study

Parul Chopra[1], Ginni Bharti[2], Ankita Gupta[1], Shyam Prakash[2], Sudip Kumar Datta[2*]

[1]*Department of Laboratory Medicine, Delhi State Cancer Institute, New Delhi*
[2]*Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi*

Article Info

[*]***Corresponding Author:***
Sudip Kumar Datta
Professor & Head,
Department of Laboratory Medicine
All India Institute of Medical Sciences
New Delhi, India – 110029
Off: 011-26594629 Mob: (O) 9868397278
E-mail: sudip.datta@aiims.edu

Keywords

Electrolyte interference, Direct vs indirect ISE, Hyperproteinemia, Hypertriglyceridemia, Pseudohyponatremia

Abstract

Background: Electrolyte analysers use ion selective electrode (ISE) technologies, either 'direct'(dISE) or 'indirect'(iISE). Indirect methods are susceptible to errors in samples with high protein or lipid content. This study evaluated the individual and combined effects of elevated total proteins(TP) and triglycerides(TG) on sodium(Na^+) and potassium(K^+) measurements.

Methods: Fifty serum samples were pooled to prepare a low-protein, low-lipid matrix. Three enrichment series were generated using graded additions of 20% human albumin, 20% Intralipid, or both to achieve varying TP and TG. All samples were analysed for Na^+ and K^+ using dISE and iISE. The difference between methods (iISE–dISE) was correlated with TP and TG using regression and Bland–Altman analyses. A multivariable model assessed predictors of electrolyte bias.

Results: Albumin addition produced negative sodium bias, reaching upto 10% at TP>13g/dL, while potassium changes remained small. Triglycerides caused marked sodium underestimation at high levels (>2000 mg/dL) with minimal effect on potassium levels. When both interferents were elevated, sodium disagreement was high (bias −5.4mmol/L; limits of agreement −19.1 to +8.2), reflecting a strong cumulative effect. Multivariable regression showed TP as the strongest predictor of Na^+ bias ($R^2 = 0.67$), with minimal contribution by TG.

Conclusion: This experimental study demonstrates that increasing concentrations of total protein and triglycerides, particularly when both are elevated produce significant negative biases in sodium measured by iISE, while potassium values remain comparable. These findings become clinically relevant in critical care/ cancer patients where they receive albumin and/or lipid infusions and dISE/ iISE results are used interchangeably for patient monitoring.

Introduction

Electrolytes like sodium (Na^+), potassium (K^+), Calcium (Ca^{++}) and Chloride (Cl^-) affect neuronal excitability, cardiac conduction, fluid balance, and acid–base homeostasis. Accurate measurement is therefore essential across critical care, oncology, renal and hepatic disease, endocrine disorders, and perioperative settings, where even small deviations can alter clinical decisions [1-5].

Electrolyte analysers use ion selective electrode (ISE) technology in laboratories and point-of-care (POC) settings. ISEs may be direct (dISE) measuring undiluted samples, or indirect (iISE), where samples are diluted before measurement. Both rely on the same electrochemical principle but differ in sample handling, matrix interaction, and susceptibility to interference. dISE, used in most arterial blood gas analysers, POC Electrolyte analysers and some dry chemistry analysers (Vitros, Ortho Clinical Diagnostics, USA), measures electrolyte activity in aqueous phase of plasma, while iISE, common on high-throughput auto-analysers, measures activity in diluted samples with dilution ratios from 1:5 to 1:46 allowing lower sample volumes. iISE assumes a plasma water fraction of approximately 93% and uses it to expand the measurable concentration range [6,7]. In hyperproteinemia or hyperlipidemia, reduced plasma water fraction leads to falsely low electrolyte values by iISE (electrolyte exclusion effect). dISE remains unaffected because no dilution is performed [8]. These discrepancies may become clinically significant in marked biochemical derangements, such as multiple myeloma, macroglobulinemia, nephrotic syndrome, severe hypertriglyceridemia, parenteral nutrition, pancreatitis, and critical illness, where precise electrolyte values are crucial for therapeutic decisions [9-11]. Thus interpretation becomes difficult when dISE and iISE are used interchangeably for follow-up.

Although the effect of high protein or lipid concentrations on electrolyte measurement is well recognised [12-14], their relative and combined influence on dISE and iISE has not been well characterised. Data are limited when both abnormalities coexist, a scenario common in patients receiving lipid infusions, albumin therapy, chemotherapy, targeted therapy, or those with complex metabolic profiles.

Therefore, this experimental study was designed to simulate and quantify the effect of increasing total protein and triglyceride concentrations—individually and in combination—on Na^+ and K^+ measured by direct and indirect ISE. Using graded enrichment series and comparing measured with calculated values, we aimed to characterise matrix effects, assess method agreement, and identify conditions producing clinically meaningful deviations.

Aims and Objectives

1. To evaluate the effect of increasing total protein on the difference between direct and indirect ISE measurements of sodium and potassium.
2. To assess the impact of increasing triglycerides on the difference between direct and indirect ISE measurements.
3. To analyse the combined influence of high protein and high triglyceride levels on sodium and potassium discrepancies.
4. To develop a multivariable regression model to predict the combined effect of total protein and triglycerides on sodium and potassium measurement bias.

Methods

Study Design and Setting

This experimental study was carried out in the Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi. Fifty leftover serum samples received for electrolyte investigations were identified based on low baseline concentrations of total protein (<6 g/dL) and triglycerides (<100 mg/dL), pooled, aliquoted and used for all our further experiments. The study protocol was approved by the Institutional Ethics Committee, All India Institute of Medical Sciences (AIIMS), New Delhi (Ref. No. IEC-789/07.10.2022). Supplementary Figure 1 depicts the study design.

Preparation of Enriched Samples

To generate a gradient of protein and lipid concentrations, 20% human albumin (AlbuRel, Reliance Life Sciences) and 20% Intralipid (Fresenius Kabi) were used. Intralipid is a sterile intravenous fat emulsion used as parenteral nutrition. These were added to aliquots of pooled serum in measured volumes (Supplementary Table 1) to generate samples with varying biochemical characteristics. Three enriched sample sets were prepared:

- Albumin-enriched (A-series): increasing total protein concentrations achieved using 20% human albumin
- Lipid-enriched (L-series): varying concentrations of 20% Intralipid added to increase triglyceride levels
- Combined (N-series): both 20% Albumin and 20% Intralipid added to achieve low, intermediate, and high protein-triglyceride combinations.

Biochemical Measurements

Each enriched sample was analysed for sodium and potassium levels using both dISE on the Ortho VITROS 4600 dry chemistry system (Ortho-Clinical Diagnostics Inc.) and iISE on Roche Cobas 8000 modular analyser (Roche Diagnostics International Ltd., Switzerland). Total protein, albumin, and triglyceride concentrations were measured on the Cobas analyser using manufacturer-approved reagents. The analytical measurable ranges of total protein (TP) and triglyceride (TG) were 0.2–12 g/dL and 8.5–885 mg/dL respectively. Since experimental enrichments often exceeded these linear ranges, calculated TP and TG concentrations were derived from the known volumes of pooled serum, albumin solution, and lipid emulsion added to each sample using mass-balance calculations based on their assumed concentrations. These calculated concentrations were used as the reference values for

downstream analysis (Supplementary Table 1).

Quality Control and Pre-Analytical Validation

Prior to sample testing, internal quality control (Levey–Jennings charts) and EQAS reports were reviewed. All analytes were within acceptable quality limits (±2 SD) on both platforms, indicating satisfactory analytical performance and readiness for the experimental procedures. To confirm comparability between instruments, two levels of SMART Lab controls were analysed on both the Cobas and VITROS platforms for Na^+ and K^+. The control results demonstrated minimal (<5%) inter-instrument variation, confirming readiness for the experimental procedures.

Data Processing and Analytical Approach

All data analyses were performed in Python (version 3.11.7) using Jupyter Notebook (version 7.0.8). NumPy (version 1.26.4) and Pandas (version 2.1.4) were used for numerical handling and array operations, while visualisations were generated using Matplotlib (version 3.8.0). Statistical analysis (paired t-tests, linear regression, Bland–Altman analysis, and multivariable modelling) were performed using SciPy (version 1.11.4) and scikit-learn (version 1.2.2).

Paired t-test were conducted for each enrichment series (A, L, and N) to compare sodium and potassium values between dISE and iISE. These tests assessed whether deviations between methods were statistically significant within each series. Linear regression assessed the agreement between measured and calculated TP and TG levels within each series. For each sample, the difference between direct and indirect ISE measurements (iISE – dISE) was calculated separately for sodium and potassium. These differences were analysed against calculated TP and TG concentrations. Agreement between dISE and iISE for Na^+ and K^+ was further evaluated using Bland–Altman analysis, yielding mean bias and 95% limits of agreement.

To visualise combined effect of TP and TG on measurement bias, two-dimensional scatter plots were created with calculated TP on the x-axis and calculated TG on the y-axis. For each sample, the size of the marker reflected the magnitude of the difference (iISE–dISE), allowing an intuitive picture of how strongly both variables contributed to interference. To make these patterns clinically meaningful, tolerance limits of ±4.0 mmol/L for sodium and ±0.3 mmol/L for potassium were applied, in accordance with CLIA allowable total error limits and biological variation–based quality specifications, as deviations beyond these thresholds are considered analytically and clinically significant [15-16].

Finally, a multivariable linear regression model was constructed using calculated TP and TG to predict measurement bias for sodium and potassium. This model evaluated whether matrix composition could predict iISE–dISE deviation and support the development of correction algorithms for samples with suspected interference.

Results

Measured sodium (Na^+) and potassium (K^+) values by direct and indirect ISE across A-, L-, and N-series are shown in Table 1. Across samples, TP ranged from 5.6–13.9 g/dL in the albumin-added A-series and 2.8–13.9 g/dL in the combined N-series, while remaining 5.7–6.3 g/dL in the lipid-only L-series. TG ranged from 105–2437 mg/dL in the L-series and 56–3602 mg/dL in the N-series, whereas A-series samples contained only baseline TG levels of 52–93 mg/dL.

Table 1: Summary of Sodium and Potassium Measurements by Direct and Indirect Ion Selective Electrodes Across Albumin-, Lipid-, and Combined-Enriched Samples.

Series	Number of samples	Na iISE (mmol/L); Mean±SD	Na dISE (mmol/L); Mean±SD	K iISE (mmol/L); Mean±SD	K dISE (mmol/L); Mean±SD	Na bias (iISE-dISE); Median (Range)	K bias (iISE-dISE); Median (Range)	Percent Na bias [(iISE-dISE)/ dISE] *100; Median (Range)	Percent K bias [(iISE-dISE)/ dISE] *100; Median (Range)
A series	17	119.54±12.4	125.21±7.8	3.3±0.8	3.37±0.8	-6.6 (-11.7, 1.2)	-0.1 (-0.2, 0.1)	-5.3 (-10.17, 0.9)	-3 (-8.3, 2.7)
L series	21	112.29±13.9	113.19±11.3	3.8±0.4	3.84±0.4	1.4 (-17.3, 2.4)	0 (-0.65, 0.15)	1.15 (-15.7, 1.9)	0 (-17.1, 3.4)
N series	33	113.56±18.9	118.98±23.5	2.41±0.6	2.56±0.6	-3.95 (-21.7, 3.25)	-0.1 (-0.4, 0)	-3.7 (-13.2, 2.8)	-4.35 (-16, 0)

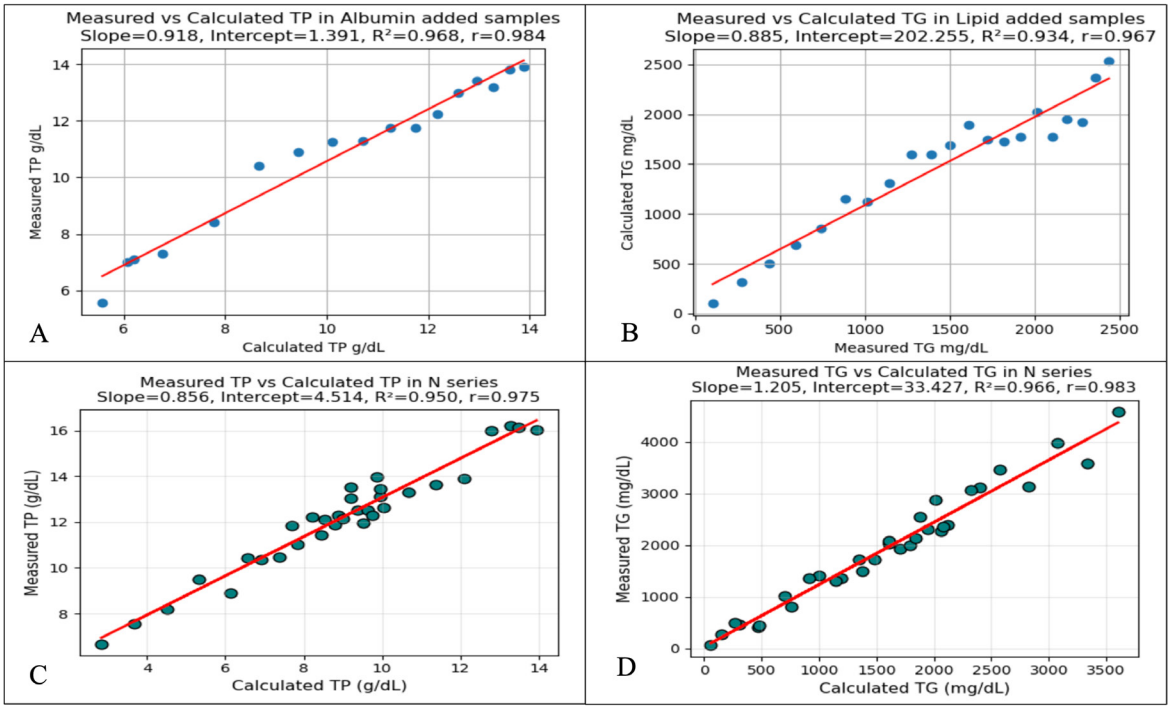
Na- Sodium; K – Potassium; dISE-Direct Ion Selective Electrodes; iISE- Indirect Ion Selective Electrodes

Data are presented as mean ± SD for measured concentrations and median (range) for differences.

Validation of Enrichment Model: Correlation Between Measured and Calculated Concentrations

Across all three enrichment series (albumin-only, lipid-only, and combined), the measured concentrations of TP and TG correlated closely with their calculated target values (Figure 1A–D). Albumin-enriched samples displayed an excellent correlation for TP ($R^2 = 0.968$), and lipid-enriched samples demonstrated similarly high correlation for TG ($R^2 = 0.934$). In the combined enrichment series, both TP and TG continued to show strong correlation ($R^2 = 0.950$ and $R^2 = 0.966$, respectively). These findings confirmed stable and predictable enrichment gradient and justify the use of calculated values for subsequent analyses.

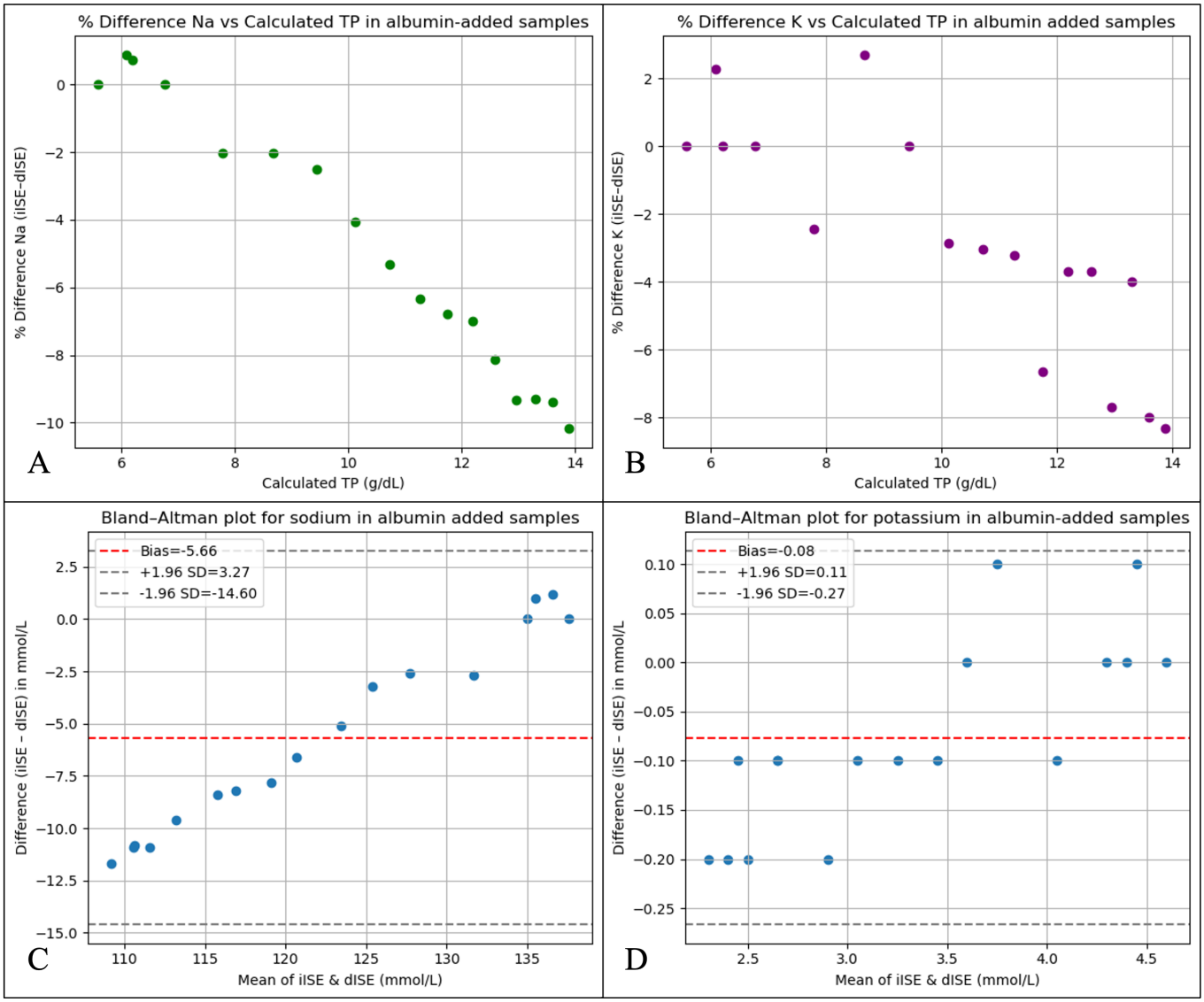
Figure 1: Linear regression analysis between measured and calculated concentrations for (A) TP in albumin-enriched samples, (B) TG in lipid-enriched samples, (C) TP in combined albumin and lipid enriched samples, and (D) TG in combined albumin + lipid enriched samples.



Effect of Albumin Addition on Sodium and Potassium Measurements (Figure 2)

In albumin-enriched samples, increasing TP levels produced progressive widening of the sodium gap between indirect and direct ISE measurements. As TP increased, sodium showed increasingly negative deviation (iISE – dISE), reaching nearly 10% underestimation at the highest protein concentrations (Figure 2A). Potassium showed a smaller downward trend (Figure 2B). Paired t-tests confirmed significant differences for both electrolytes, with sodium showing a mean bias of −5.665 mmol/L ($p < 0.05$) and potassium a smaller but significant mean bias of −0.076 mmol/L ($p < 0.05$). Bland–Altman plots supported these findings with sodium demonstrating a marked negative bias with wide limits of agreement (Figure 2C), whereas potassium remained comparatively stable, exhibiting minimal bias and tight limits (Figure 2D). Collectively, these results highlight the strong susceptibility of sodium, more than potassium, to protein-related analytical interference.

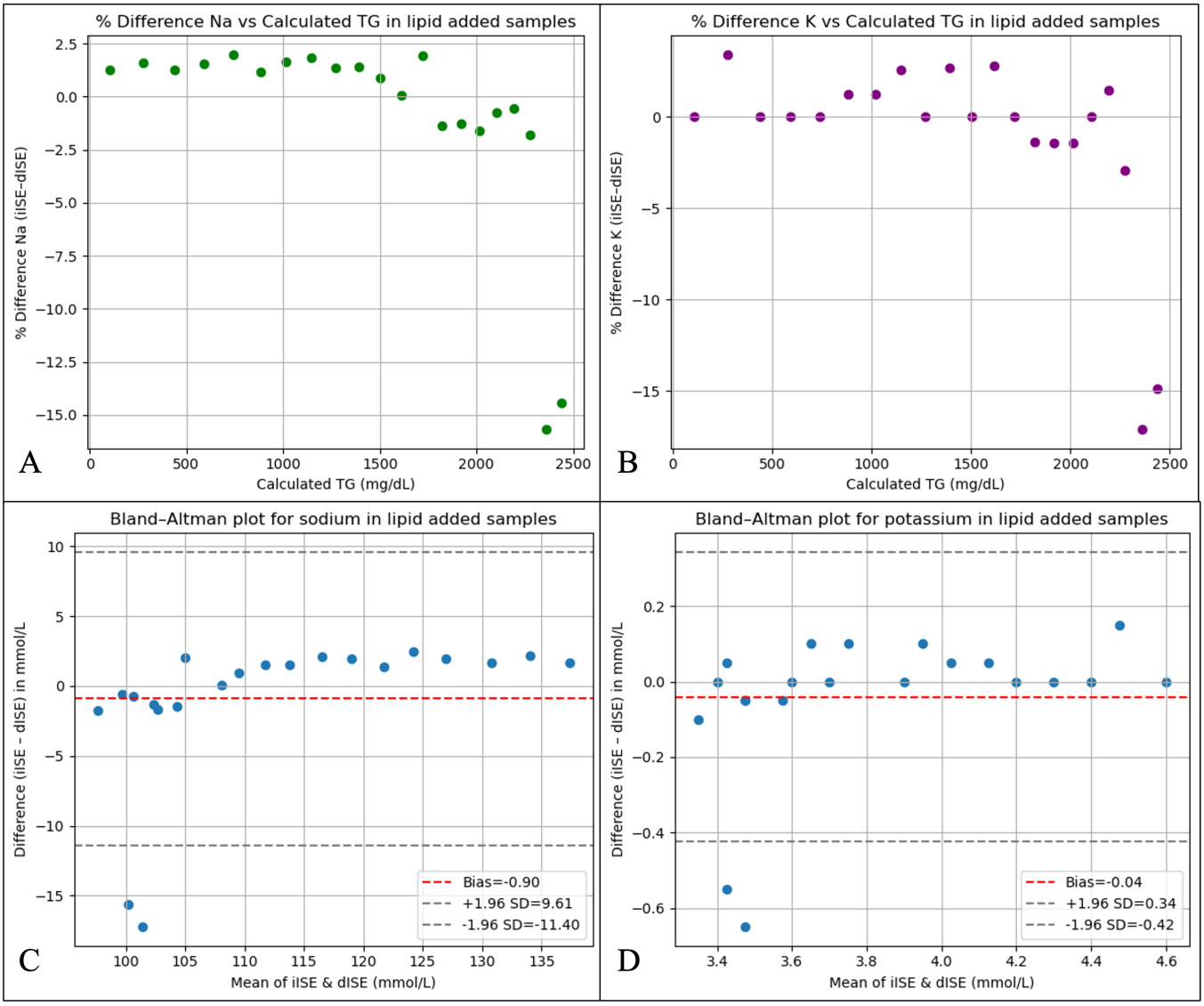
Figure 2: Effect of albumin addition on sodium and potassium measurements.
(A) %Difference in sodium (iISE–dISE) across calculated total protein levels.
(B) %Difference in potassium (iISE–dISE) across calculated total protein levels.
(C) Bland–Altman plot for sodium (iISE vs dISE) showing bias and 95% limits of agreement.
(D) Bland–Altman plot for potassium (iISE vs dISE) with bias and limits of agreement.



Effect of Triglyceride Addition on Sodium and Potassium Measurements

In the triglyceride-enriched samples, sodium values remained close to baseline but began to show a negative shift once triglyceride concentrations exceeded 2000 mg/dL (Figure 3A). Potassium values displayed only small and gradual changes with rising TG (Figure 3B). Paired t-tests showed that neither sodium (mean difference −0.895 mmol/L, p = 0.4529) nor potassium (mean difference −0.040 mmol/L, p = 0.3552) exhibited statistically significant deviation between indirect and direct ISE, indicating lipids alone had a weaker impact on measurement bias. Bland–Altman analysis supported this observation, demonstrating only mild sodium bias and negligible potassium bias across the lipid-added series (Figure 3C–D).

Figure 3: Effect of triglyceride addition on sodium and potassium measurements.
(A) %Difference in sodium (iISE–dISE) across calculated triglyceride levels in lipid-added samples.
(B) %Difference in potassium (iISE–dISE) across calculated triglyceride levels in lipid-added samples.
(C) Bland–Altman plot for sodium (iISE vs dISE) showing bias and 95% limits of agreement.
(D) Bland–Altman plot for potassium (iISE vs dISE) with bias and limits of agreement.

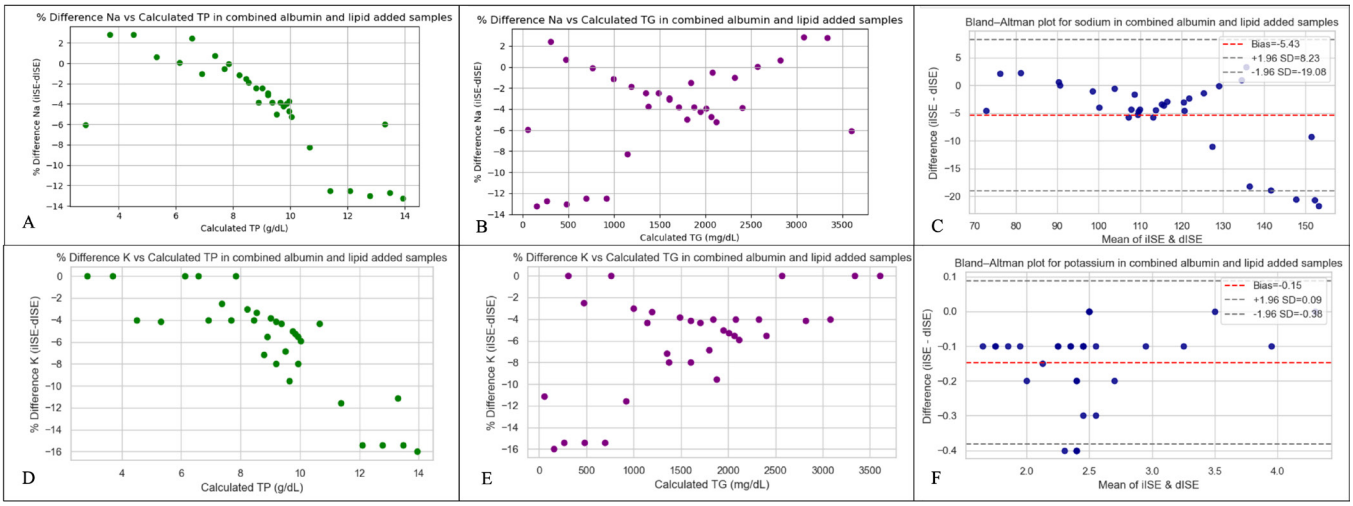


Combined Effect of Albumin and Triglyceride Addition

In samples where both TP and TG were simultaneously elevated, the impact on electrolyte measurements became more pronounced. Sodium showed progressively larger negative deviations between indirect and direct ISE measurements as both TG and TP concentrations increased (Figure 4A, 4B). The greatest underestimation of sodium was observed at the highest TP and TG levels, indicating an amplified interference. Bland–Altman analysis for sodium demonstrated a substantial negative bias, with wide limits of agreement, reflecting poor concordance between direct and indirect ISE measurements under dual interference conditions (Figure 4C). This widening of limits was greater than that observed in either albumin-only or lipid-only enrichment series.

Potassium showed increased deviation in the combined enrichment model, with negative differences becoming more evident at higher TP and TG concentrations (Figure 4D, 4E). However, the magnitude of potassium deviation remained smaller compared to sodium. Bland–Altman analysis for potassium showed a modest negative bias with relatively narrow limits of agreement, indicating better stability despite combined interference (Figure 4F).

Figure 4: Effect of combined albumin and triglyceride enrichment on sodium and potassium measurements by indirect and direct ISE.
(A) % Difference in sodium (iISE–dISE) across calculated total protein concentrations
(B) % Difference in sodium (iISE–dISE) across calculated total triglyceride concentrations
(C) Bland–Altman plot for sodium showing the mean bias and 95% limits of agreement
(D) % Difference in potassium (iISE–dISE) plotted against calculated protein concentrations
(E) % Difference in potassium (iISE–dISE) plotted against calculated triglyceride concentrations
(F) Bland–Altman plot for potassium illustrating the mean bias and 95% limits of agreement

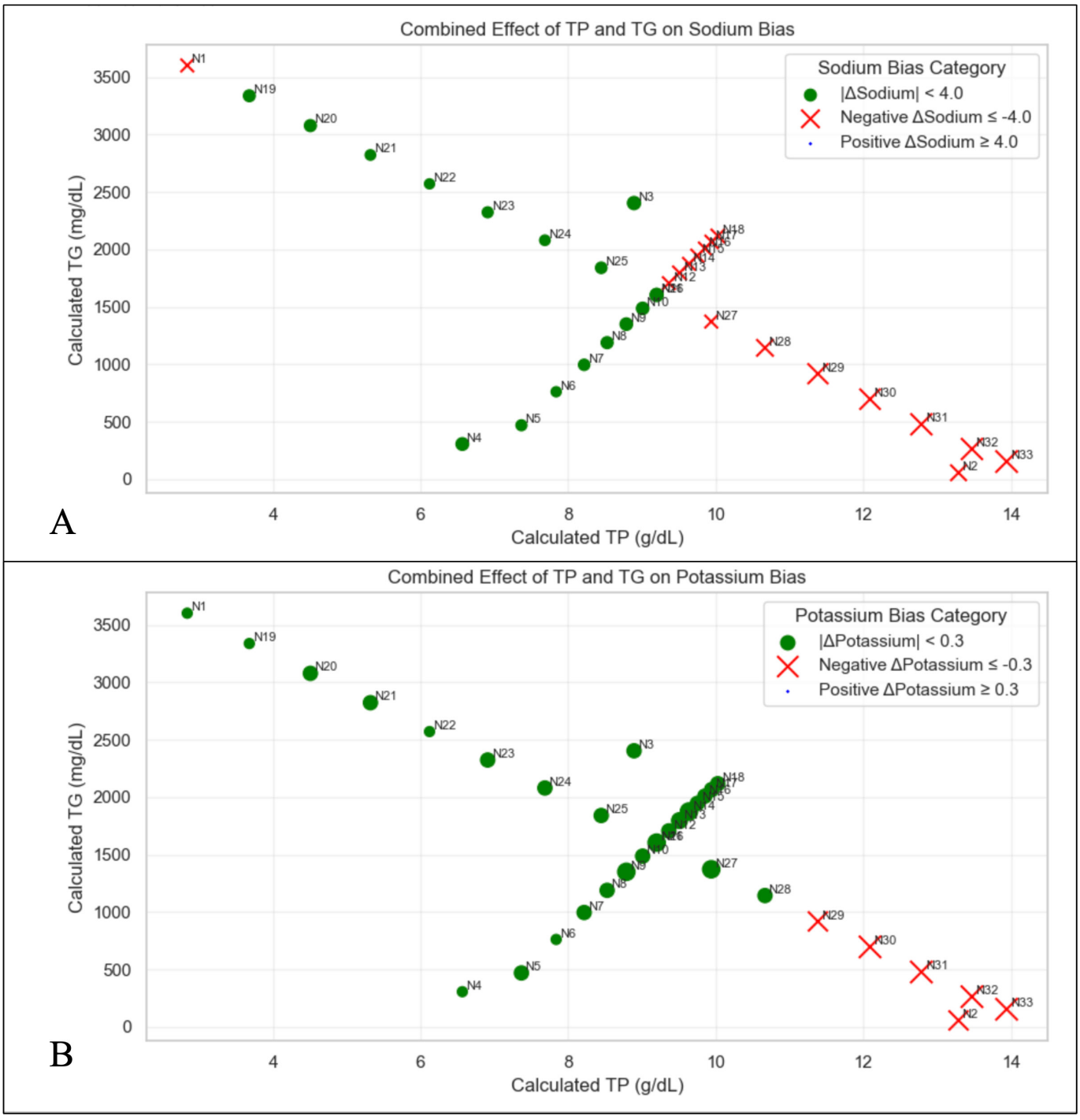


Statistical comparison confirmed that both electrolytes exhibited significant differences between indirect and direct ISE in the combined series, with sodium showing a mean difference of −5.426 mmol/L ($p < 0.05$) and potassium a mean difference of −0.147 mmol/L ($p < 0.05$), supporting cumulative interference effect.

Integrated Influence of Total Protein and Triglycerides

To assess combined effect of TP and TG on electrolyte measurement bias, calculated TP and TG concentrations were plotted simultaneously in a two-dimensional space, with marker size proportional to the magnitude of the difference between indirect and direct ISE measurements (iISE–dISE). For sodium, a clear coordinated interference pattern was observed (Figure 5A).

Figure 5: Combined effect of total protein and triglyceride levels on sodium and potassium measurement differences.
(A) Scatter plot illustrating the combined influence of calculated total protein and triglyceride concentrations on sodium bias (iISE–dISE), categorized using a clinical cut-off of ±4.0 mmol/L.
(B) Scatter plot showing the combined influence of calculated total protein and triglyceride concentrations on potassium bias (iISE–dISE), categorized using a clinical cut-off of ±0.3 mmol/L.



**The size of each marker reflects the magnitude of deviation, with larger markers indicating greater measurement discrepancies.*

Samples with concurrent elevations of TP and TG demonstrated the largest negative sodium deviations, with multiple points exceeding the clinically relevant cut-off of ±4.0 mmol/L. These high-magnitude deviations clustered predominantly at the upper extremes of both TP and TG, indicating a strong additive effect of proteins and lipids on sodium underestimation by indirect ISE.

In contrast, potassium showed substantially greater analytical stability (Figure 5B). When a clinical threshold of ±0.3 mmol/L was applied, most samples remained within acceptable limits, even at higher TP and TG concentrations. Only a small subset of samples with very high TP values demonstrated potassium deviations approaching or exceeding this cut-off, and the overall magnitude of deviation was markedly smaller than that observed for sodium.

Increasing marker size across both plots visually highlights that clinically meaningful electrolyte discrepancies arise primarily under conditions of combined biochemical elevation, with sodium being far more susceptible than potassium. These findings reinforce that simultaneous hyperproteinemia and hypertriglyceridemia pose the greatest risk for significant sodium misclassification when measured by indirect ISE, whereas potassium measurements remain comparatively valid.

Multivariable Linear Regression for Predicting Sodium and Potassium Bias

Multivariable regression analysis was performed on the combined-matrix (N-series) samples to quantify the effect of TP and TG on electrolyte deviation. For sodium, the regression model demonstrated good explanatory performance (R^2 = 0.6655) with a mean absolute error (MAE) of 3.19 mmol/L.

Diff_Na (iISE − dISE)= −2.4383(TP) − 0.0012(TG) + 18.2809

From the model, a corrective equation for estimating direct ISE sodium from indirect ISE measurements is:

Na dISE-equivalent =

Na(iISE)+2.4383(TP)+0.0012(TG)−18.2809

For potassium, the model also showed moderate predictive strength (R^2 = 0.6652) with a very small MAE of 0.059 mmol/L.

Diff_K (iISE − dISE) = −0.03733(TP) − 0.0(TG) + 0.18966

Based on the multivariable regression model, a corrective equation for estimating direct ISE potassium from indirect ISE measurements is:

K dISE-equivalent =K (iISE)+0.03733(TP)−0.18966

Across both electrolytes, TP was the strongest driver of deviation, whereas triglycerides contributed modestly to sodium bias and had no measurable impact on potassium. These models indicate that errors can be predicted using TP and TG levels. However, external validation is required before clinical use.

Discussion

This experimental study evaluated how progressively increasing TP and TG concentrations added individually and in combination affect the differences between direct and indirect ion-selective electrode (ISE) measurements for sodium and potassium. By experimenting with a graded and controlled enrichment model that mimicked the biochemical extremes seen in real clinical practice, matrix effects responsible for ISE discrepancies were characterised. A key strength of the study was the use of a pooled, low interference serum matrix with stepwise enrichment with albumin and lipid. Strong correlation between measured and calculated TP and TG concentrations across the A-, L-, and N-series confirmed analytical validity of the model. These findings give confidence that the experimental matrix accurately represented progressive elevations similar to those seen in real clinical scenarios. Clinically extreme TP and TG levels occur in conditions such as multiple myeloma, Waldenström macroglobulinemia, severe dehydration, and certain inflammatory states, familial chylomicronemia, poorly controlled diabetes, pancreatitis, and patients receiving parenteral nutrition.

The results from albumin-enriched samples align with prior studies demonstrating pseudohyponatremia in hyperproteinemia, with underestimation becoming more pronounced at high protein concentrations ~>8 g/dL. Dimeski and colleagues, 2012 [12] showed similar behaviour using paraprotein-rich clinical samples, and Fortgens & Pillay, 2011[14] emphasised that this effect persists even with modern analysers. Potassium showed only minimal change, consistent with prior literature suggesting that potassium measurement is less susceptible to water-displacement error due to its smaller absolute concentration and tighter physiological buffering. Paired t-test comparing sodium and potassium values between dISE and iISE gave significant p-values ($p < 0.05$ for both sodium and potassium). This supports earlier clinical observations published by Chopra & Datta, 2020 [17] in their direct–indirect ISE comparison study.

In the triglyceride-enriched samples, the results were consistent with prior evidence of lipid-related underestimation of sodium by iISE [18-20]. Koch et al. 2021 [18] demonstrated that Intralipid®-based testing underestimates lipid interference, with minimal sodium bias up to lipemic indices approaching 2000, whereas patient samples showed significant pseudohyponatremia (≥4 mmol/L) at lipemic indices exceeding ~700, reflecting poor correlation between lipemic index and true interference. They also reported a poor correlation between lipemic index and triglyceride concentration. Sen, 2016 [19] similarly showed a progressive negative bias in iISE sodium levels with increasing TG concentration, while potassium remained comparatively stable. Hou et al, 2025 [20] confirmed concentration-dependent sodium bias with rising TG concentrations using controlled lipemia models where potassium deviations remained small and clinically insignificant. These findings align closely with our lipid-enriched series, where sodium bias increased at higher triglyceride concentrations while potassium remained relatively robust. Notably, in our lipid-only series, differences between indirect and direct ISE were not statistically significant

for either sodium or potassium indicating that moderate hyperlipidemia alone may not consistently result in measurable analytical disagreement.

The combined enrichment model (N-series) demonstrated how simultaneous elevations of TP and TG amplify electrolyte measurement bias beyond the effect of either interferent alone (Figure 4A–C). Bland–Altman analysis also showed negative sodium bias with markedly widened limits of agreement, indicating poor agreement under combined interference conditions. On the other hand, potassium had increased deviation in the combined series as compared to albumin-only or lipid-only models. These findings extend prior observations of pseudohyponatremia in hyperproteinemia or hyperlipidemia studied in isolation by showing a clear cumulative and potentially synergistic electrolyte exclusion effect when both matrix components are elevated together.

In the combined enrichment model, TP was the dominant driver of bias, particularly for sodium. Although rising TG levels did contribute to increasing sodium deviation, the largest negative biases were seen when high TG concentrations occurred together with elevated TP. This suggests that proteins are the main driver of the electrolyte exclusion effect, with lipids acting to further intensify the error rather than causing it independently.

Potassium deviations increased in the combined series than in the A or L series, but the overall magnitude remained small. Only a few samples exceeded a clinically relevant cut-off of ±0.3 mmol/L. Sodium, however, frequently crossed the ±4.0 mmol/L allowable error limit, underscoring its greater susceptibility to exclusion effects when measured by iISE. Overall, TP is the primary determinant of clinically significant sodium bias with TG acting as an amplifier. This difference in susceptibility helps explain why pseudohyponatremia is commonly encountered in complex biochemical states, whereas clinically meaningful potassium errors are far less frequent.

With a two-dimensional visual framework incorporating the clinical thresholds, this study provides a practical method for identifying biochemical conditions under which iISE results may become diagnostically misleading. To our knowledge, this visual–analytical approach to evaluating combined matrix interference has not been previously reported and represents a novel methodological contribution to the study of electrolyte measurement interference. These findings emphasise the need for heightened laboratory vigilance and support preferential use of dISE or confirmatory testing in patients with complex biochemical abnormalities. Multivariable regression analysis of the combined enrichment (N-series) samples showed how TP and TG together influence electrolyte bias. TP emerged as the primary driver of sodium deviation, while TG contributed modestly and did not influence potassium. This behaviour is in line with the electrolyte exclusion mechanism described by Dimeski et al. 2012 [12], who identified displacement of plasma water by non-aqueous solids causes pseudohyponatremia, while potassium remains relatively resistant to clinically meaningful error. Recent modelling studies have similarly shown that sodium bias can be estimated from matrix parameters, supporting algorithm-based interpretation or automated flagging in high-risk samples [12,20]. Our dual-interference model demonstrates that combined protein and lipid measurements can predict sodium bias under complex biochemical conditions, supporting multivariable approaches to improve interpretation of indirect ISE results in routine practice. The most novel and clinically relevant aspect of this study is the evaluation of the combined enrichment model (N-series), which remains largely unexplored. Earlier studies typically examined hyperproteinemia or hyperlipidemia in isolation; however, many real-world settings - such as nephrotic syndrome, myeloma, Waldenström macroglobulinemia, parenteral nutrition, chemotherapy-related metabolic derangements, and critical illness - feature simultaneous elevations of both proteins and lipids. In this context, our controlled enrichment model provides important new insight into their combined effects. The findings are clinically relevant because sodium measurement directly influences management in hyponatremia. Misclassification may lead to inappropriate hypertonic saline therapy, inadequate treatment, or risk of osmotic demyelination if true sodium levels are higher than reported. There are, however, limitations to consider. Although the enrichment model mimics clinical conditions, added albumin and Intralipid do not perfectly replicate endogenous proteins and lipoproteins. The study used a single direct and single indirect ISE platform; thus, results may differ with other analyser architectures. Furthermore, extreme concentrations achieved experimentally, particularly triglycerides >3000 mg/dL, may not occur in all practice settings. Finally, regression models derived from artificial samples require validation in real patient populations before clinical use. Overall, this study provides an experimental evaluation of how proteins and lipids - alone and together - affect sodium and potassium measurement by direct and indirect ISE. The findings highlight the importance of recognising matrix effects when interpreting electrolytes in complex biochemical states and support the use of direct ISE or blood-gas-based technologies in high-interference scenarios. Future clinical studies should validate these correction models and explore automated laboratory warning systems for samples at risk of pseudohyponatremia.

Conclusion

This study shows that increases in TP and TG, especially when they occur together, can substantially distort sodium results measured by indirect ISE, while direct ISE measurements remain largely reliable. TP was the main driver of sodium error, with triglycerides adding further bias only at very high levels. When both were elevated, the largest discrepancies were observed. Potassium measurements were far less affected across all conditions. These findings highlight the need for

caution when interpreting indirect ISE sodium in patients with complex biochemical abnormalities and support the use of direct ISE or confirmatory testing to prevent clinically meaningful errors.

Funding
This research received no external funding.

Conflicts of Interest
The authors declare that they have no competing interests.

Author Contributions
Conceptualization: SKD, PC; Sample collection, experiments: GB, PC; Formal analysis: PC; Supervision: SKD, SP; Writing original draft: PC, AG; Writing, review & editing: PC, AG, GB, SKD,SP; Approval of final version: PC, AG, GB, SKD,SP
All authors have read and approved the final manuscript.

Acknowledgements
The authors acknowledge the technical staff of the Clinical Biochemistry and SMART Laboratory, Department of Laboratory Medicine, AIIMS New Delhi for their support in selection and processing of samples.

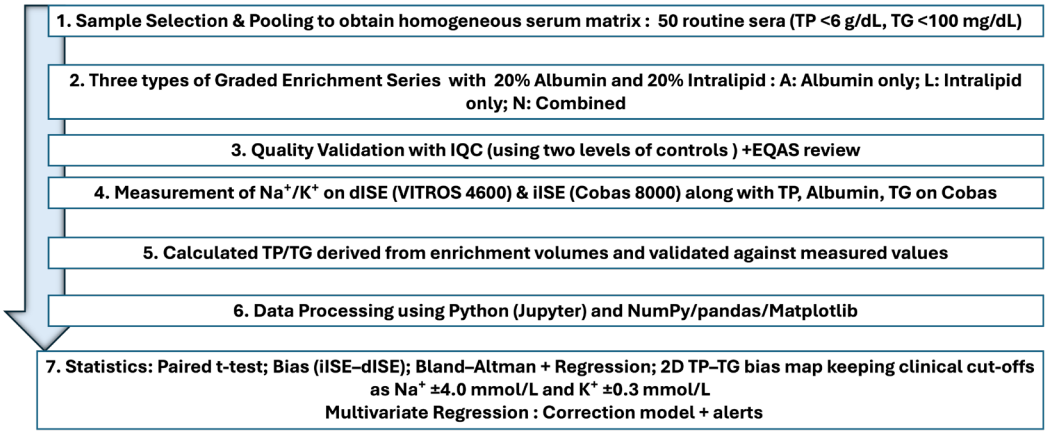
References

1. Bachmann KF, Hess B, Koitmäe M, Bloch A, Regli A, Reintam Blaser A. Electrolyte disorders in the critically ill: a retrospective analysis. Sci Rep. 2025;15:13943. doi: 10.1038/s41598-025-98677-7.
2. Pinter J, Canaud B, Mayne KJ, Stuard S, Moissl U, Kooman J, Jager KJ, Chesnaye NC, Smyth B, Genser B. Associations of Abnormal Fluid Status, Plasma Sodium Disorders, and Low Dialysate Sodium with Mortality in Patients on Hemodialysis. Clin J Am Soc Nephrol. 2024;19:1444-1452. doi: 10.2215/CJN.0000000000000552.
3. Papper S. Fluid and electrolyte disturbances in cirrhosis. Am J Med Sci. 1976;272:53-6. doi: 10.1097/00000441-197607000-00006.
4. Fries C, Fenske W. Elektrolytentgleisungen unter endokrinologischen Gesichtspunkten [An Endocrinological Perspective on Electrolyte Imbalances]. Dtsch Med Wochenschr. 2025;150:883-889. doi: 10.1055/a-2318-7580.
5. Hasselgren E, Hertzberg D, Camderman T, Björne H, Salehi S. Perioperative fluid balance and major postoperative complications in surgery for advanced epithelial ovarian cancer. Gynecol Oncol. 2021;161:402-407. doi: 10.1016/j.ygyno.2021.02.034.
6. Dimeski G, Badrick T, John AS. Ion Selective Electrodes (ISEs) and interferences--a review. Clin Chim Acta. 2010;411:309-17. doi: 10.1016/j.cca.2009.12.005.
7. Datta SK, Chopra P. Interference in Ion-Selective Electrodes Due to Proteins and Lipids. J Appl Lab Med. 2022;7:589-595. doi: 10.1093/jalm/jfab125.
8. Weld BA, Morgan TJ, Presneill JJ, Weier S, Cowley D. Plasma sodium measurements by direct ion selective methods in laboratory and point of care may not be clinically interchangeable. J Clin Monit Comput. 2017;31:1103-1109. doi: 10.1007/s10877-016-9938-1.
9. Ladenson JH, Apple FS, Koch DD. Misleading hyponatremia due to hyperlipemia: a method-dependent error. Ann Intern Med. 1981;95:707-8. doi: 10.7326/0003-4819-95-6-707.
10. Ladenson JH, Apple FS, Aguanno JJ, Koch DD. Sodium measurements in multiple myeloma: two techniques compared. Clin Chem. 1982;28:2383-2386.
11. Steinberger BA, Ford SM, Coleman TA. Intravenous immunoglobulin therapy results in post-infusional hyperproteinemia, increased serum viscosity, and pseudohyponatremia. Am J Hematol. 2003;73:97-100. doi: 10.1002/ajh.10325.
12. Dimeski G, Morgan TJ, Presneill JJ, Venkatesh B. Disagreement between ion selective electrode direct and indirect sodium measurements: estimation of the problem in a tertiary referral hospital. J Crit Care. 2012;27:326.e9-16. doi: 10.1016/j.jcrc.2011.11.003.
13. Aziz F, Sam R, Lew SQ, Massie L, Misra M, Roumelioti ME, Argyropoulos CP, Ing TS, Tzamaloukas AH. Pseudohyponatremia: Mechanism, Diagnosis, Clinical Associations and Management. J Clin Med. 2023;12:4076. doi: 10.3390/jcm12124076.
14. Fortgens P, Pillay TS. Pseudohyponatremia revisited: a modern-day pitfall. Arch Pathol Lab Med. 2011;135:516-9. doi: 10.5858/2010-0018-RS.1.
15. CLIA. Clinical Laboratory Improvement Amendments of 1988 (CLIA). Westgard QC Quality Requirements: 2025 CLIA Acceptance Limits for Proficiency Testing [Internet]; 2025. Available from: https://westgard.com/clia-a-quality/quality-requirements/2024-clia-requirements.html
16. Centers for Medicare & Medicaid Services (CMS). Clinical Laboratory Improvement Amendments (CLIA): Proficiency testing acceptance limits. Code of Federal Regulations, Title 42, Part 493. Washington (DC): U.S. Government Publishing Office; updated 2024. Available from: https://www.ecfr.gov/current/title-42/chapter-IV/subchapter-G/part-493/subpart-I
17. Chopra P, Datta SK. Discrepancies in Electrolyte Measurements by Direct and Indirect Ion Selective Electrodes due to Interferences by Proteins and Lipids. J Lab Physicians. 2020;12:84-91. doi: 10.1055/s-0040-1713690.
18. Koch CD, Vera MA, Messina J, Price N, Durant TJS, El-Khoury JM. Preventing pseudohyponatremia: Intralipid®-based lipemia cutoffs for sodium are inappropriate. Clin Chim Acta. 2021;520:63-66. doi: 10.1016/j.cca.2021.05.032.
19. Sen S. A study on effect of lipemia on electrolyte

measurement by direct ion selective electrode method. J Biomol Res Ther. 2016;5:142.
20. Hou Y, Liao J, Wang Z, Xie C. Mitigating electrolyte measurement discrepancies in high triglycerides samples: a comparative analysis of direct and indirect ISE for sodium, potassium and chloride measurements. Scand J Clin Lab Invest. 2025;85:599-607. doi: 10.1080/00365513.2025.2585468.

Supplementary files

Supplementary Figure 1: Overview of the study workflow and analytical strategy.



Supplementary Table 1: Composition of experimentally enriched serum samples and calculation of theoretical total protein (TP) and triglyceride (TG) concentrations used for validation experiments.

Supplementary Table 1A. A series: Samples spiked with 20% human albumin							
Label	Pooled serum added (µL)	Amount of albumin added (µL)	Measured TP levels (g/dL)	Mass of TP (g) in serum (600µL)	Mass of TP(g) in Albumin added	Total TP Mass (g)	Calculated TP (g/dL)
A1	600	0	5.58	0.03	0.00	0.03	5.58
A2	600	20	7.00	0.03	0.00	0.04	6.08
A3	600	25	7.11	0.03	0.01	0.04	6.20
A4	600	50	7.30	0.03	0.01	0.04	6.77
A5	600	100	8.40	0.03	0.02	0.05	7.78
A6	600	150	10.40	0.03	0.03	0.06	8.66
A7	600	200	10.90	0.03	0.04	0.08	9.44
A8	600	250	11.25	0.03	0.05	0.09	10.12
A9	600	300	11.30	0.03	0.06	0.10	10.72
A10	600	350	11.75	0.03	0.07	0.11	11.26
A11	600	400	11.75	0.03	0.08	0.12	11.75
A12	600	450	12.25	0.03	0.09	0.13	12.19
A13	600	500	13.00	0.03	0.11	0.14	12.59
A14	600	550	13.40	0.03	0.12	0.15	12.95
A15	600	600	13.20	0.03	0.13	0.16	13.29
A16	600	650	13.80	0.03	0.14	0.17	13.60
A17	600	700	13.90	0.03	0.15	0.18	13.88

TP mass added from albumin was calculated assuming a TP concentration of 21 g/dL in the albumin preparation, using the formula (21 000 / 100 000) × albumin volume (µL). Total TP mass was used to derive the calculated TP concentration after volume correction.

Supplementary Table 1B. L series: Samples spiked with 20% Intralipid							
Label	**Pooled serum added (µL)**	**Amount of LIPID (1:5 dilution of 20% Intralipid) added (µL)**	**Measured TG levels (mg/dL)**	**Mass of TG (mg) in Serum (800µL)**	**Mass of TG (mg) in LIPID added**	**Total TG Mass (mg)**	**Calculated TG levels (mg/dL)**
L1	800	0	104.85	0.84	0	0.84	104.85
L2	800	20	316.1	0.84	1.42	2.26	275.46
L3	800	40	500.1	0.84	2.84	3.68	437.95
L4	800	60	686.25	0.84	4.26	5.10	592.88
L5	800	80	857.45	0.84	5.68	6.52	740.77
L6	800	100	1146.65	0.84	7.1	7.94	882.09
L7	800	120	1126.4	0.84	8.52	9.36	1017.26
L8	800	140	1308.1	0.84	9.94	10.78	1146.68
L9	800	160	1598.3	0.84	11.36	12.20	1270.71
L10	800	180	1596.5	0.84	12.78	13.62	1389.67
L11	800	200	1688.2	0.84	14.2	15.04	1503.88
L12	800	220	1895.2	0.84	15.62	16.46	1613.61
L13	800	240	1749.05	0.84	17.04	17.88	1719.12
L14	800	260	1730.25	0.84	18.46	19.30	1820.64
L15	800	280	1769.45	0.84	19.88	20.72	1918.41
L16	800	300	2025.8	0.84	21.3	22.14	2012.62
L17	800	320	1775.25	0.84	22.72	23.56	2103.46
L18	800	340	1945.35	0.84	24.14	24.98	2191.12
L19	800	360	1920.6	0.84	25.56	26.40	2275.76
L20	800	380	2369.55	0.84	26.98	27.82	2357.53
L21	800	400	2534.8	0.84	28.4	29.24	2436.57

TG mass added from lipid was calculated assuming a TG concentration of 7100 mg/dL in the lipid preparation, using the formula (7100/100000) × lipid volume (µL). Total TG mass was used to derive calculated TG concentration after volume correction.

Supplementary Table 1C: N series: Samples spiked with both 20% human albumin and 20% Intralipid.

Label	Pooled serum added (µL)	Amount of albumin added (µL)	Amount of LIPID (1:5 dilution of 20% Intralipid) added (µL)	Measured TP levels (g/dL)	Measured TG levels (mg/dL)	Mass of TP (g) in serum	Mass of TP (g) in albumin added	Mass of TP (g) in LIPID added	Total TP Mass (g)	Calculated TP (g/dL)	Mass of TG (mg) in Serum	Mass of TG (mg) in LIPID added	Mass of TG (mg) in Albumin added	Total TG Mass (mg)	Calculated TG levels (mg/dL)
N1	600.00	0.00	600.00	6.67	4587.10	0.03	0.00	0.00	0.03	2.84	0.63	42.60	0.00	43.23	3602.43
N2	600.00	600.00	0.00	16.21	75.00	0.03	0.13	0.00	0.16	13.29	0.63	0.00	0.05	0.68	56.43
N3	600.00	600.00	600.00	12.28	3122.90	0.03	0.13	0.00	0.16	8.89	0.63	42.60	0.05	43.28	2404.28
N4	600.00	50.00	20.00	10.43	467.60	0.03	0.01	0.00	0.04	6.57	0.63	1.42	0.00	2.05	306.43
N5	600.00	100.00	40.00	10.47	418.60	0.03	0.02	0.00	0.05	7.37	0.63	2.84	0.01	3.48	469.88
N6	600.00	150.00	80.00	11.02	813.20	0.03	0.03	0.00	0.07	7.84	0.63	5.68	0.01	6.32	761.58
N7	600.00	200.00	120.00	12.23	1408.70	0.03	0.04	0.00	0.08	8.22	0.63	8.52	0.02	9.17	996.21
N8	600.00	250.00	160.00	12.10	1353.30	0.03	0.05	0.00	0.09	8.53	0.63	11.36	0.02	12.01	1189.02
N9	600.00	300.00	200.00	11.87	1718.50	0.03	0.06	0.00	0.10	8.79	0.63	14.20	0.02	14.85	1350.28
N10	600.00	350.00	240.00	12.15	1714.80	0.03	0.07	0.00	0.11	9.01	0.63	17.04	0.03	17.70	1487.15
N11	600.00	400.00	280.00	13.53	2032.70	0.03	0.08	0.00	0.12	9.20	0.63	19.88	0.03	20.54	1604.77
N12	600.00	450.00	320.00	12.53	1936.90	0.03	0.09	0.00	0.13	9.36	0.63	22.72	0.04	23.39	1706.94
N13	600.00	500.00	360.00	11.94	1992.40	0.03	0.11	0.00	0.14	9.51	0.63	25.56	0.04	26.23	1796.51
N14	600.00	550.00	400.00	12.53	2545.60	0.03	0.12	0.00	0.15	9.64	0.63	28.40	0.04	29.07	1875.68
N15	600.00	600.00	440.00	12.28	2305.90	0.03	0.13	0.00	0.16	9.75	0.63	31.24	0.05	31.92	1946.16
N16	600.00	650.00	480.00	13.98	2878.80	0.03	0.14	0.00	0.17	9.85	0.63	34.08	0.05	34.76	2009.31
N17	600.00	700.00	520.00	13.13	2274.40	0.03	0.15	0.00	0.18	9.95	0.63	36.92	0.06	37.61	2066.21
N18	600.00	750.00	560.00	12.63	2403.40	0.03	0.16	0.00	0.19	10.03	0.63	39.76	0.06	40.45	2117.75
N19	600.00	50.00	560.00	7.57	3575.90	0.03	0.01	0.00	0.04	3.68	0.63	39.76	0.00	40.39	3338.27
N20	600.00	100.00	520.00	8.17	3989.40	0.03	0.02	0.00	0.06	4.51	0.63	36.92	0.01	37.56	3078.45
N21	600.00	150.00	480.00	9.49	3138.60	0.03	0.03	0.00	0.07	5.32	0.63	34.08	0.01	34.72	2822.85
N22	600.00	200.00	440.00	8.89	3472.10	0.03	0.04	0.00	0.08	6.12	0.63	31.24	0.02	31.89	2571.38
N23	600.00	250.00	400.00	10.35	3063.00	0.03	0.05	0.00	0.09	6.91	0.63	28.40	0.02	29.05	2323.93
N24	600.00	300.00	360.00	11.84	2367.40	0.03	0.06	0.00	0.10	7.69	0.63	25.56	0.02	26.21	2080.40
N25	600.00	350.00	320.00	11.42	2137.90	0.03	0.07	0.00	0.11	8.45	0.63	22.72	0.03	23.38	1840.72
N26	600.00	400.00	280.00	13.04	2088.20	0.03	0.08	0.00	0.12	9.20	0.63	19.88	0.03	20.54	1604.77
N27	600.00	450.00	240.00	13.44	1503.30	0.03	0.09	0.00	0.13	9.94	0.63	17.04	0.04	17.71	1372.49
N28	600.00	500.00	200.00	13.29	1315.70	0.03	0.11	0.00	0.14	10.67	0.63	14.20	0.04	14.87	1143.78
N29	600.00	550.00	160.00	13.64	1359.00	0.03	0.12	0.00	0.15	11.38	0.63	11.36	0.04	12.03	918.56
N30	600.00	600.00	120.00	13.91	1020.00	0.03	0.13	0.00	0.16	12.09	0.63	8.52	0.05	9.20	696.75
N31	600.00	650.00	80.00	15.99	442.60	0.03	0.14	0.00	0.17	12.79	0.63	5.68	0.05	6.36	478.28
N32	600.00	700.00	40.00	16.15	507.10	0.03	0.15	0.00	0.18	13.47	0.63	2.84	0.06	3.53	263.07
N33	600.00	750.00	20.00	16.03	271.50	0.03	0.16	0.00	0.19	13.94	0.63	1.42	0.06	2.11	153.95

For the N-series, TP contributed by each component was first converted to mass using the component volume (µL) and assumed TP concentration (serum 5.58 g/dL; albumin 21 g/dL; lipid 0.1 g/dL) as: TP mass = volume × concentration / 100000. Total TP mass was obtained by summing component masses, and calculated TP (g/dL) was derived as: (total TP mass × 100000) / total volume (µL).

For the N-series, TG mass contributed by each component was calculated from its volume (µL) and assumed TG concentration (serum 104.85 mg/dL; lipid 7100 mg/dL; albumin 8 mg/dL) using TG mass (mg) = concentration × volume / 100000. Total TG mass was obtained by summing component masses, and calculated TG (mg/dL) was derived as (total TG mass / total volume) × 100000.

Review Article

Association of Circulating and Urinary Calprotectin (S100A8/A9, MRP8/14) Levels with Type 2 Diabetes Mellitus and its Complications: A Systematic Review and Meta-analysis

Roshan Kumar Mahat[1*], Mritunjay Kumar Mishra[2], Nand Kishor Prasad Sah[3], Sushil Yadav[1]

[1]*Department of Biochemistry, Teerthanker Mahaveer Medical College & Research Centre, Teerthanker Mahaveer University, Moradabad, India*

[2]*Department of Biochemistry, Faculty of Medicine and Health Sciences (FMHS), SGT University, Gurugram, Haryana, India*

[3]*Department of Physiotherapy, Teerthanker Mahaveer University, Moradabad, India*

Article Info

[*]***Corresponding Author:***
Roshan Kumar Mahat, MSc (Med), PhD (Med), MBA (HHM)
Professor of Biochemistry
Teerthanker Mahaveer Medical College & Research Centre, Teerthanker Mahaveer University, Moradabad, India
E-mail: mahatroshan79@gmail.com
ORCID: 0000-0003-1202-6227

Keywords

Calprotectin, Type 2 diabetes mellitus, Diabetic complications, Meta-analysis

Abstract

Background: Calprotectin is a myeloid-derived damage-associated molecular pattern implicated in metabolic inflammation and vascular injury. Evidence regarding circulating and urinary calprotectin in type 2 diabetes mellitus (T2DM) and its complications remains inconsistent.

Methods: PubMed, Scopus, and Web of Science were searched up to 25th October 2025. Human observational studies reporting quantitative calprotectin levels were included. Standardized mean differences (SMDs) with 95% confidence intervals (CIs) were calculated using a random-effects model. Heterogeneity was assessed using I^2 and Cochran's Q. Sensitivity analyses, prediction intervals, and publication bias evaluation using the Doi plot and LFK index were performed.

Results: Seventeen studies were included in the meta-analysis. Circulating calprotectin levels were significantly higher in patients with T2DM compared with healthy controls (SMD = 1.08, 95% CI: 0.77–1.39; $p < 0.0001$; $I^2 = 78.2\%$). Sensitivity analyses confirmed the robustness of this association, although small-study effects were suggested. Higher circulating calprotectin levels were observed in T2DM patients with peripheral neuropathy (SMD = 1.02; 95% CI: 0.06–1.98) and macrovascular complications (SMD = 0.43; 95% CI: 0.08–0.77) compared with those without complications, but these associations showed considerable heterogeneity and limited stability. No significant difference in circulating calprotectin levels was observed between T2DM patients with microalbuminuria and normoalbuminuria. Urinary calprotectin levels did not differ significantly between T2DM patients and controls.

Conclusions: Circulating calprotectin is elevated in T2DM, supporting its role as a biomarker of metabolic inflammation. Associations with neuropathy and macrovascular disease are suggestive but heterogeneous, while evidence for urinary calprotectin remains inconclusive.

Introduction

Diabetes mellitus (DM) represents a significant global public health challenge and ranks among the most prevalent non-communicable diseases worldwide. Current estimates indicate that approximately 537 million adults, accounting for 10.5% of the global population, are affected by diabetes, with projections suggesting an increase to 784 million by the year 2045 [1]. Type 2 diabetes mellitus (T2DM), the predominant form of the disease, is characterized by substantial heterogeneity in clinical presentation and disease progression. This variability often results in delayed diagnosis, multiple pathophysiological abnormalities, and differing susceptibilities to complications. These complications are broadly classified into microvascular complications, which include retinopathy, neuropathy, and nephropathy, and macrovascular complications, such as cardiovascular, cerebrovascular, and peripheral vascular diseases [2].

Calprotectin, also known as myeloid-related protein 8/14 (MRP8/14), is a stable heterodimer constituted of calcium-binding proteins MRP8 (Calgranulin A, S100A8) and MRP14 (Calgranulin B, S100A9) [3]. MRP8 and MRP14 belonging to the S100 protein family, are predominantly expressed in cells of myeloid lineage, particularly in monocytes/macrophages and neutrophils [4] with additional expression observed in platelets [5]. MRP8/14 exhibits both intracellular and extracellular functions. Intracellularly, it regulates calcium homeostasis, interacts with cytoskeletal and microtubule components, and contributes to intracellular trafficking within phagocytes [6]. It's essential role in leukocyte transmigration has recently been demonstrated in murine models [7]. When released, calprotectin functions as a damage-associated pattern molecules (DAMP), promoting the inflammatory response [6]. Calprotectin is actively secreted during phagocyte stress responses [8] and has long been acknowledged as a robust biomarker of inflammation [9]. Recent evidence identifies calprotectin as an endogenous activator of Toll-like receptor 4 (TLR4) and a ligand for the receptor for advanced glycation end products (RAGE) [10, 11], both of which play pivotal roles in the pathophysiological pathways that drive microvascular and macrovascular complications in T2DM.

Although numerous individual studies have explored circulating and urinary calprotectin levels in patients with T2DM and its associated complications, the existing evidence remains fragmented, and no comprehensive pooled analysis has been conducted to date. Consequently, this systematic review and meta-analysis aimed to synthesize the available data to elucidate the association between calprotectin and T2DM along with its complications. Specifically, we aimed to compare circulating and urinary calprotectin levels between individuals with T2DM and healthy controls, as well as to examine whether calprotectin concentrations differ between T2DM patients with complications and those without complications.

Materials and Methods

This meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [12] and was prospectively registered with PROSPERO under the registration number CRD420251176512.

Data sources and search strategy

A comprehensive literature search was conducted in MEDLINE/PubMed, Scopus, and Web of Science, encompassing all studies published up to 25 October 2025. The search strategy utilized a combination of keywords and medical subject headings, including type 2 diabetes mellitus, diabetes mellitus, type 2 diabetes, diabetes, diabetics, type 2 DM, T2DM, T2D, calprotectin, myeloid-related protein 8/14, MRP 8/14, S100A8/A9 and leukocyte protein L1. The complete search strategy is detailed in Supplementary Table 1. Furthermore, the reference lists of eligible articles and pertinent review papers were manually screened to identify additional studies. Subsequent to the database searches, two independent reviewers (RKM and MKM) initially screened the titles and abstracts to exclude studies that did not meet the inclusion criteria. The full texts of potentially eligible articles were then assessed in detail to confirm their suitability for inclusion in the systematic review. Any disagreements between the reviewers were resolved through discussion among the authors until a consensus was achieved.

Inclusion and Exclusion criteria

Study selection was guided by the Population, Exposure, Comparator, Outcomes, and Study Design (PECOS) framework. Population (P): Adults diagnosed with type 2 diabetes mellitus (T2DM), with or without diabetes-related complications. Exposure (E): Quantitative measurements of circulating (serum or plasma) or urinary calprotectin levels (S100A8/A9 or MRP8/14). Comparator (C): Healthy, non-diabetic individuals or subjects with normal glucose tolerance (NGT); and/or T2DM patients without complications for analyses involving complication subgroups. Outcomes (O): Differences in circulating calprotectin levels between T2DM patients and healthy controls, and between T2DM patients with and without complications; differences in urinary calprotectin levels among these groups; and comparisons across specific diabetic complications when reported. Study Design (S): Observational human studies, including cross-sectional, case–control, and cohort studies, as well as interventional studies that provided baseline quantitative calprotectin data.

Exclusion criteria encompassed studies involving type 1 diabetes, gestational diabetes, animal or in vitro models, studies lacking an appropriate control or comparator group, and those that did not include quantitative calprotectin measurements (e.g., qualitative assessments, reviews, case reports, editorials). Furthermore, duplicate publications or overlapping datasets, conference abstracts lacking accessible full texts, and studies

with inadequate extractable data were also omitted from this review.

Data extraction

Data from each included study were extracted according to predefined criteria, which included the following elements: the first author's surname, year of publication, country of study, study design, study population, diagnostic criteria for diabetes, sample characteristics, study quality, and key findings. In instances where critical information was missing or ambiguous, the corresponding authors were contacted for clarification. Numerical data presented solely in graphical format were extracted utilizing WebPlotDigitizer (Version 5). The data extraction process was carried out independently by two reviewers, with all extracted information subsequently verified by a third reviewer to ensure accuracy. Any discrepancies identified during the extraction process were resolved through discussion and consensus among the reviewers.

Quality assessment

The methodological quality of case–control and prospective cohort studies was assessed using the Newcastle–Ottawa Scale (NOS) [13], while cross-sectional studies were evaluated using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Analytical Cross-Sectional Studies [14]. For studies assessed using the NOS, a maximum of 9 points could be awarded, and study quality was categorized as low quality (≤4 points), moderate quality (5–6 points), or high quality (≥7 points). For studies assessed using the JBI checklist, quality was determined based on the proportion of affirmative ("yes") responses. Studies achieving ≤49% "yes" responses were classified as having a high risk of bias (low quality), those with 50–69% were considered to have a moderate risk of bias (moderate quality), and those with ≥70% were deemed to have a low risk of bias (high quality). Two reviewers (RKM and MKM) independently assessed the methodological quality of the included studies. A third reviewer (NKPS) subsequently verified the assessments, and any discrepancies were resolved through consensus following discussion among all reviewers.

Data synthesis and analysis

Data analyses were performed using R software version 4.4.1 (R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria). Statistical procedures were conducted with the "meta" package (version 8.0-2) and "metasens" package (version 1.5-2). For studies presenting calprotectin data as medians with interquartile ranges (IQRs), corresponding means and standard deviations (SDs) were estimated using the Meta-analysis Accelerator [15]. In the study by Pedersen et al. [16], calprotectin values in healthy controls were reported as medians with 2.5th–97.5th percentiles due to non-normal distribution; because no additional dispersion metrics were available, the arithmetic mean was approximated by using the median as the mean, and an SD was estimated under the assumption of approximate normality. These converted values were used solely for quantitative pooling. Pooled estimates were expressed as standardized mean differences (SMDs) with 95% confidence intervals (CIs). A random-effects model was employed to calculate overall effect sizes. Forest plots and drapery plots were generated to visually summarize pooled analyses. Additionally, prediction intervals (PIs) were computed to quantify the expected dispersion of true effect sizes across future studies. Heterogeneity was assessed using Cochran's Q test and the inconsistency statistic (I^2), where $I^2 > 50\%$ indicated substantial heterogeneity. A p value < 0.10 for Cochran's Q test was considered evidence of significant heterogeneity. Given the presence of notable heterogeneity, leave-one-out sensitivity analyses were conducted to evaluate the influence of individual studies on the pooled estimate. Potential publication bias was assessed using doi plots, and asymmetry was quantified using the Luis Furuya-Kanamori (LFK) index [17], where values within ±1 indicated no asymmetry, between ±1 and ±2 suggested minor asymmetry, and > ±2 denoted major asymmetry. Subgroup analyses were performed to explore sources of heterogeneity based on diagnostic criteria (ADA, WHO, not reported), country, sample size (≥100 vs <100), and type of biological sample (serum vs plasma). For Pedersen et al. [16], plasma values from T2DM patients were classified under the plasma subgroup. Although controls were measured in serum, the study demonstrated no significant difference between serum and plasma calprotectin concentrations (p = 0.191), therefore it was appropriate to classify the comparison under the plasma subgroup. With the exception of Cochran's Q test, p-values of 0.05 or less were considered statistically significant.

Results

Characteristics of included studies

The literature search across major databases identified 171 records from PubMed/MEDLINE, 219 from Scopus, and 205 from Web of Science, yielding a total of 595 records. After removal of 250 duplicate records, 345 unique records remained for screening. During the title and abstract screening phase, 288 records were excluded, leaving 57 articles for full-text retrieval and assessment. All 57 full-text articles were successfully retrieved and assessed for eligibility. Of these, 43 studies were excluded for the following reasons: publication in a non-English language (n = 4); measurement of calprotectin in biological samples other than blood or urine, or absence of circulating calprotectin data (n = 8); studies conducted exclusively in pregnant women (n = 1); conference abstracts or oral presentations without accessible full text (n = 6); editorials (n = 1); absence of calprotectin data (n = 8); studies not specific to type 2 diabetes mellitus (n = 5); separate measurement of S100A8 and S100A9 rather than the MRP8/14 complex (n = 3); lack of group-wise calprotectin data (n = 5); and data not extractable for quantitative synthesis (n = 2). An additional 3

records were identified through citation searching. A total of 17 studies met the eligibility criteria and were included in the final systematic review and meta-analysis [16, 18-33]. The details of the search process are shown in Figure 1. The study characteristics as well as qualities of the included studies are shown in Table 1.

Figure 1: PRISMA flow diagram for the study selection process.

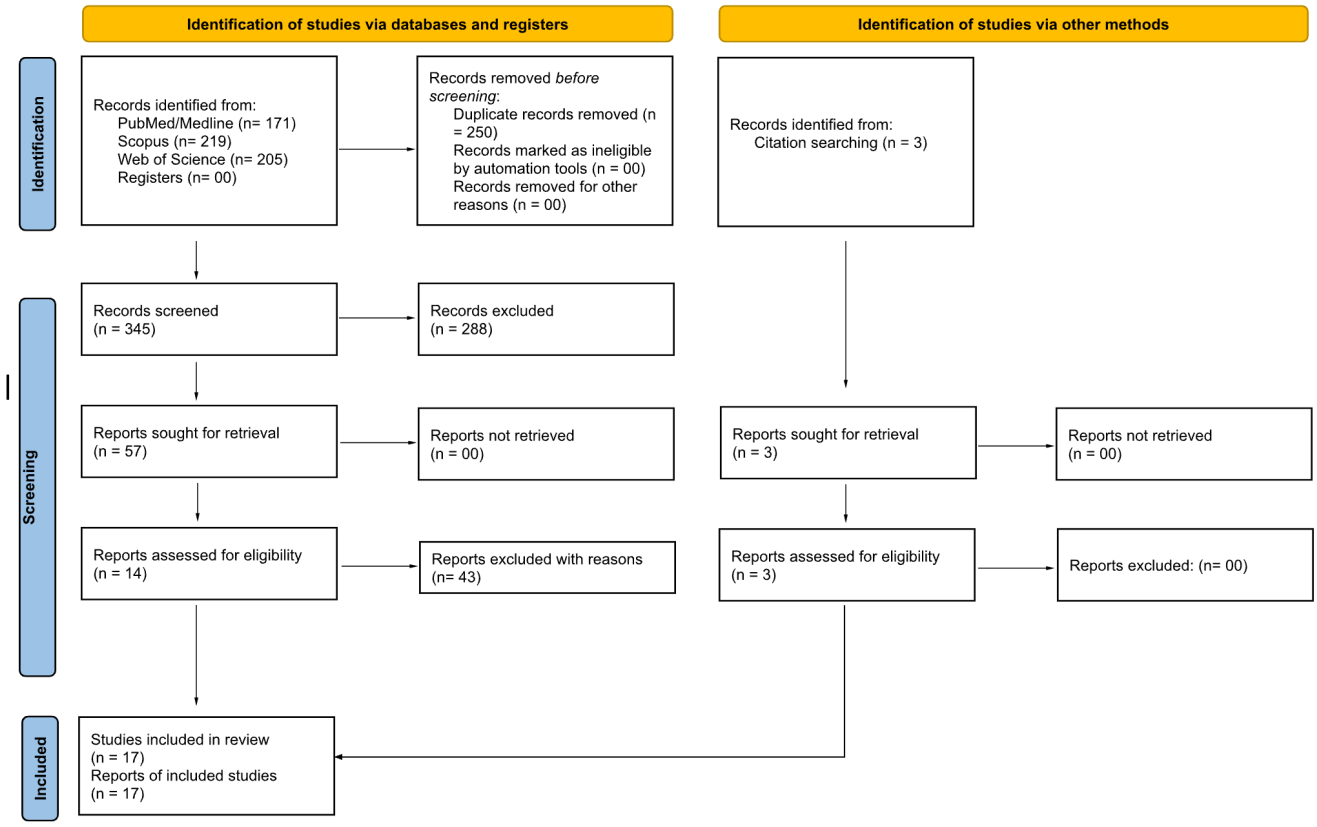


Table 1: Baseline characteristics of studies included in meta-analysis.

Author	Year	Country	Study design	Study subjects	Diagnostic criteria of type 2 DM	Sample	Quality	Main findings
Afify et al. [18]	2022	Egypt	C-C	20 healthy controls, 20 T2DM without diabetic peripheral neuropathy (No-DPN), 20 T2DM with neuropathy (DPN)	NR	Serum	High	Serum calprotectin levels were significantly elevated in patients with T2DM compared to control subjects. Moreover, these levels were found to be highest in T2DM patients with diabetic peripheral neuropathy (DPN) in comparison to those without neuropathy.
El-adawy et al. [19]	2017	Egypt	C-C	20 healthy controls and 20 T2DM	ADA	Serum	High	Calprotectin levels were significantly elevated in individuals with T2DM compared to control subjects.
El-hafez et al. [20]	2021	Egypt	C-C	15 controls, 15 T2DM without neuropathy, 60 T2DM with neuropathy	NR	Serum	Moderate	Calprotectin levels were significantly elevated in individuals with diabetic neuropathy compared to those without neuropathy and control subjects.

Author	Year	Country	Study design	Study subjects	Diagnostic criteria of type 2 DM	Sample	Quality	Main findings
Gao et al. [21]	2020	China	C-S	35 T2DM and 43 controls	WHO	Serum	High	Serum calprotectin levels in T2DM patients were significantly higher than those in healthy controls.
Gobejishvili et al. [22]	2024	USA	C-C	17 T2DM adults and 11 healthy controls	NR	Serum	Moderate	Calprotectin levels are elevated in individuals with T2DM when compared to non-diabetic controls; however, this difference is statistically insignificant.
Gu et al. [23]	2020	China	C-S	80 T2DM (20 normoalbuminuria, 20 microalbuminuria, 20 macroalbuminuria with normal serum creatinine and 20 macroalbuminuria with increased serum creatinine)	NR	Serum	High	As diabetic kidney disease progressed, there was a gradual increase in the serum levels of MRP8/14.
Halawa et al. [24]	2020	Egypt	C-C	60 T2DM (30 with PAD & 30 without PAD) and 30 controls	NR	Plasma and urine	High	Both plasma and urinary calprotectin levels were significantly elevated in individuals with T2DM compared to control subjects. Additionally, plasma calprotectin levels were notably higher in T2DM patients with PAD. However, the urinary calprotectin levels in T2DM patients with PAD were comparable to those in T2DM patients without PAD.
Hamzah et al. [25]	2020	Iraq	C-C	60 T2DM and 60 controls	WHO	Serum	Moderate	Patients with T2DM exhibited significantly elevated levels of calprotectin in comparison to healthy control subjects.
Hassan et al. [26]	2025	Iraq	C-C	40 T2DM and 20 controls	NR	Serum	Moderate	There was a significant increase in calprotectin levels in individuals with T2DM compared to those in the healthy control group.
Lylloff et al. [27]	2017	Denmark	Prospective cohort study	18 subjects with NGT and 29 with T2DM	NR	Plasma	High	Patients with T2DM exhibit elevated plasma levels of calprotectin in comparison to individuals with normal glucose tolerance.
Murad et al. [28]	2019	Saudi Arabia	C-S	200 T2DM and 50 controls	ADA	Urine	High	Urinary calprotectin is unable to differentiate between individuals with treated-controlled T2DM and those with treated-uncontrolled T2DM, nor can it distinguish between diabetic patients and non-diabetic control subjects.
Pedersen et al. [16]	2014	Denmark	C-S	305 T2DM (183 with CVD & 122 without CVD) and 120 healthy controls	WHO	Serum (controls), Plasma (T2DM)	High	The median plasma calprotectin concentration was significantly elevated in individuals with T2DM compared to the healthy reference population; however, there was no significant difference in median plasma calprotectin concentrations between patients with and without cardiovascular disease (CVD).
Peng et al. [29]	2011	China	C-S	375 T2DM (205 with CAD and 170 without CAD)	WHO	Serum	High	Diabetic patients with coronary artery disease exhibited significantly elevated plasma levels of MRP8/14 in comparison to control subjects.

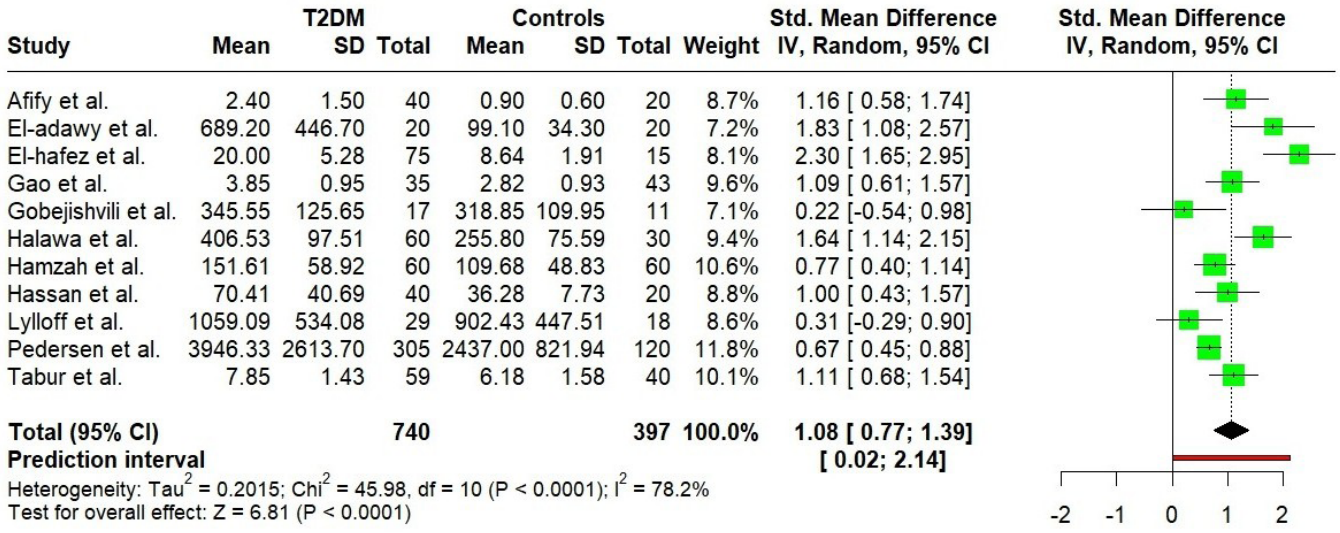
Salem et al. [30]	2024	Iraq	C-S	126 T2DM (63 with DPN and 63 without DPN)	NR	Serum	High	There were no statistically significant differences in the mean serum calprotectin levels between diabetic patients with and without DPN.
Schmaderer et al. [31]	2014	Germany	C-S	86 T2DM (46 with normoalbuminuria & 40 with microalbuminuria)	ADA	Serum	High	MRP8/14 levels were significantly elevated in the microalbuminuric cohort of patients with diabetes when compared to the normoalbuminuric cohort.
Tabur et al. [32]	2015	Turkey	C-C	59 T2DM (29 with DPN and 30 without DPN) and 40 healthy controls	NR	Serum	High	Serum calprotectin levels were significantly elevated in both patients with neuropathy and those without, compared to healthy controls. Furthermore, serum calprotectin levels were higher in diabetic patients with neuropathy than in those without neuropathy.
Velayutham et al. [33]	2021	India	C-S	126 T2DM (63 with DPN and 63 without DPN)	NR	Serum	High	The serum calprotectin levels were significantly elevated in patients with diabetic peripheral neuropathy compared to those without neuropathy.

C-C: Case-Control; C-S: Cross-Sectional; ADA: American Diabetes Association; WHO: World Health Organization; NR: Not Reported; T2DM: Type 2 Diabetes Mellitus; DPN: Diabetic Peripheral Neuropathy; PAD: Peripheral Artery Disease; CVD: Cardiovascular Disease; NGT: Normal Glucose Tolerance; CAD: Coronary Artery Disease; MRP8/14: Myeloid-related Protein 8/14.

Calprotectin in T2DM and controls

The meta-analysis of 11 studies [16, 18-22, 24-27, 32] comparing circulating calprotectin levels between individuals with T2DM (n = 740) and healthy controls (n = 397) demonstrated a significantly higher pooled standardized mean difference (SMD) among patients with T2DM (SMD = 1.08; 95% CI: 0.77–1.39; $p < 0.0001$), as shown in the overall forest plot (Figure 2).

Figure 2: Forest plot showing comparison of circulating calprotectin levels between T2DM and controls.



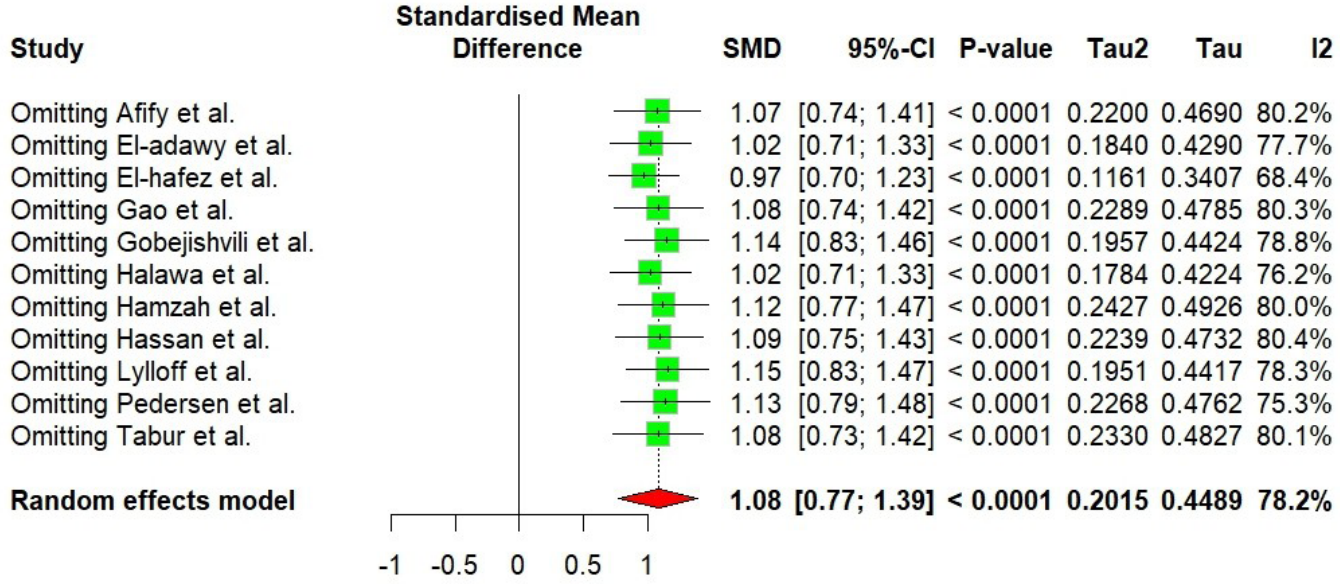
Considerable heterogeneity was present ($I^2 = 78.2\%$), and the 95% prediction interval (0.02–2.14) indicated that future studies are expected to show effects consistently favoring elevated calprotectin in T2DM, though with varying magnitude. A drapery plot illustrating the meta-analysis results using the p-value functions of individual studies is presented in Supplementary Figure 1.

Subgroup analyses provided insights into sources of heterogeneity. Stratification by country revealed the highest effect size among studies from Egypt (SMD = 1.71; 95% CI: 1.25–2.17), followed by moderate effects in other countries including China, USA, and Turkey (SMD = 0.90; 95%

CI: 0.44–1.36), Iraq (SMD = 0.84; 95% CI: 0.53–1.15), and Denmark (SMD = 0.60; 95% CI: 0.31–0.88), with a statistically significant subgroup difference (p = 0.0009) (Supplementary Figure 2). Studies with smaller sample sizes (<100) reported larger effect sizes (SMD = 1.18; 95% CI: 0.80–1.57) compared with those enrolling ≥100 participants (SMD = 0.69; 95% CI: 0.51–0.88; p = 0.0246), suggesting a small-study effect (Supplementary Figure 3). Subgrouping by biospecimen type showed that serum-based studies yielded higher pooled estimates (SMD = 1.17; 95% CI: 0.80–1.53) than plasma-based studies (SMD = 0.87; 95% CI: 0.20–1.55), although the difference was not statistically significant (p = 0.4526) (Supplementary Figure 4). Finally, subgrouping by diagnostic criteria indicated meaningful variation: studies not reporting diagnostic criteria produced higher pooled effects (SMD = 1.12; 95% CI: 0.64–1.60), whereas those adopting WHO criteria showed lower effect sizes (SMD = 0.76; 95% CI: 0.56–0.97), with a significant subgroup difference (p = 0.0156) (Supplementary Figure 5). Collectively, these figures demonstrate a consistent and significant elevation of circulating calprotectin levels in T2DM across multiple analytical layers, while highlighting potential contributors to between-study heterogeneity.

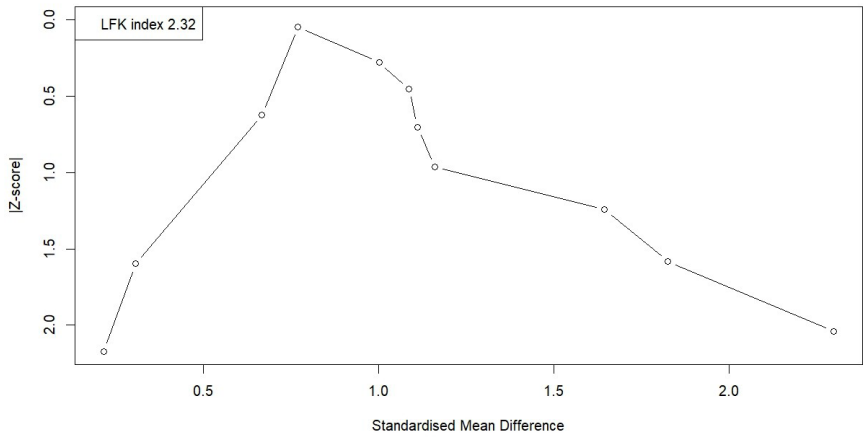
Sensitivity analysis using a leave-one-out approach demonstrated that the overall association between circulating calprotectin levels and T2DM was highly stable. Sequential omission of individual studies yielded pooled SMDs ranging from 0.97 to 1.15, with all estimates remaining statistically significant and heterogeneity values changing only minimally, indicating that no single study disproportionately influenced the overall effect (Figure 3).

Figure 3: Results of leave-one-out method in sensitivity analysis.



Assessment of potential publication bias using the DOI plot and LFK index revealed marked asymmetry (LFK = 2.32), suggesting the presence of small-study effects (Figure 4).

Figure 4: Doi plot showing publication bias.



Despite this asymmetry, the robustness of the sensitivity analysis supports the reliability of the pooled findings, reinforcing the conclusion that circulating calprotectin levels are consistently elevated in T2DM across studies.

Calprotectin levels in T2DM complications

T2DM with peripheral neuropathy

Across five studies [18, 20, 30, 32, 33] comparing individuals with and without diabetic peripheral neuropathy (DPN), circulating calprotectin levels were higher in those with DPN, with a pooled SMD of 1.02 (95% CI: 0.06–1.98; p = 0.0382). However, substantial heterogeneity was present (I^2 = 94.7%), indicating marked variability across studies (Supplementary Figure 6). The leave-one-out sensitivity analysis showed that the pooled effect was not stable, with effect sizes ranging widely from 0.69 to 1.31 and the loss of statistical significance when several studies were omitted (Supplementary Figure 7). This suggests that the overall association is sensitive to individual datasets and should be interpreted cautiously. The DOI plot demonstrated minor asymmetry (LFK = 1.64), consistent with small-study effects but not indicative of major publication bias (Supplementary Figure 8). Together, these findings support a possible association between elevated calprotectin and DPN, but highlight that the evidence is heterogeneous and not robust across analytical scenarios.

T2DM with microalbuminuria

Across two studies [23, 31] involving 60 individuals with microalbuminuria (MAU) and 66 with normoalbuminuria (NAU), calprotectin levels did not differ significantly between groups (SMD = 0.43; 95% CI: −0.06 to 0.91; p = 0.0850). Heterogeneity was low to moderate (I^2 = 41.7%) (Supplementary Figure 9). These findings indicate that circulating calprotectin does not appear to distinguish microalbuminuria from normoalbuminuria in T2DM.

T2DM with macrovascular complications

Across three studies [16, 24, 29] involving 418 individuals with macrovascular complications and 322 without, calprotectin levels were significantly higher in the macrovascular group (SMD = 0.43; 95% CI: 0.08–0.77; p = 0.0148). Moderate heterogeneity was present (I^2 = 76.6%) (Supplementary Figure 10). The leave-one-out sensitivity analysis revealed that the observed association between calprotectin and macrovascular complications in T2DM is only partially robust. Although the direction of effect remained positive in all iterations, statistical significance was lost in two of the three leave-one-out scenarios, indicating that the pooled estimate is sensitive to the inclusion of specific studies. Moderate–high heterogeneity persisted across all iterations (Supplementary Figure 11). These results suggest that while calprotectin is elevated in individuals with macrovascular complications, the strength of this association should be interpreted cautiously.

Urinary calprotectin in T2DM

Across two studies [24, 28] evaluating urinary calprotectin in T2DM (n = 260) versus controls (n = 80), the pooled effect showed no significant difference between groups (SMD = 1.19; 95% CI: −1.28 to 3.65; p = 0.3450). The studies reported highly inconsistent results, reflected in extreme heterogeneity (I^2 = 98.3%) (Supplementary Figure 12). Overall, current evidence does not demonstrate a reliable elevation of urinary calprotectin in T2DM.

Discussion

In this systematic review and meta-analysis, we conducted a comprehensive evaluation of circulating and urinary calprotectin (S100A8/A9, MRP8/14) levels in individuals diagnosed with T2DM and associated complications. The principal finding of this analysis indicates that circulating calprotectin levels are significantly elevated in patients with T2DM in comparison to healthy controls, demonstrating a large pooled effect size and consistent directionality across multiple sensitivity and subgroup analyses. These results provide strong support for the notion that calprotectin serves as an indicator of the heightened inflammatory environment characteristic of T2DM, thereby underscoring its potential utility as a biomarker for metabolic inflammation.

Calprotectin is highly expressed in activated neutrophils and monocytes, and it is released during cellular stress, inflammation, and endothelial activation. As a damage-associated molecular pattern (DAMP) molecule, it plays a significant role in amplifying innate immune responses through its interaction with Toll-like receptor 4 (TLR4) and the receptor for advanced glycation end products (RAGE) [3, 4, 6, 8-11]. Both receptors are critically involved in the pathogenesis of insulin resistance, endothelial dysfunction, and atherogenesis in T2DM [34, 35]. The markedly elevated levels of circulating calprotectin observed in patients with T2DM in our pooled analysis are, therefore, biologically plausible and align with studies that identify chronic low-grade inflammation as a defining characteristic of T2DM [36, 37]. Furthermore, emerging evidence indicates that calprotectin may play an active role in metabolic dysregulation rather than merely serving as a marker of inflammation. Increased expression of S100A8/A9 has been documented in the visceral adipose tissue of individuals with T2DM, showing a correlation with macrophage infiltration and indices of insulin resistance [38]. Consequently, elevated circulating calprotectin may function both as a biomarker and a mediator of metabolic inflammation in T2DM.

Despite the strong overall association, considerable heterogeneity was observed across the studies. Subgroup analyses identified several factors contributing to this variability, including geographic region, sample size, and the diagnostic criteria employed for T2DM. Larger effect sizes were noted in studies conducted in Egypt and within smaller cohorts, indicating potential small-study effects and population-

specific inflammatory profiles. Notably, sensitivity analyses demonstrated that the overall association remained stable and statistically significant regardless of the exclusion of individual studies, thereby reinforcing the robustness of the primary finding.

Our analysis indicates that circulating calprotectin levels are elevated in patients with T2DM who exhibit peripheral neuropathy and macrovascular complications, in comparison to those without such complications; however, these associations were not as robust as the primary comparison. The relationship with diabetic peripheral neuropathy (DPN) was marked by significant heterogeneity and sensitivity to individual studies, suggesting that the current evidence remains preliminary. Nonetheless, this finding is consistent with experimental data that implicate S100A8/A9-mediated neuroinflammation and microvascular dysfunction in the pathogenesis of diabetic neuropathy [18, 20, 32].

In our meta-analysis, we observed that patients with macrovascular complications exhibited elevated levels of calprotectin. Calprotectin is predominantly synthesized by immune cells of myeloid lineage, particularly macrophages and neutrophilic granulocytes, which may facilitate atherosclerotic plaque development through multiple mechanisms. These mechanisms include the activation of endothelial cells that line blood vessels, leading to a disruption of their structural integrity, promotion of apoptosis, and the production of reactive oxygen species. Consequently, this cascade results in oxidative stress and the oxidation of low-density lipoproteins. Furthermore, heightened levels of calprotectin in the bloodstream are believed to augment the activation of the receptor for advanced glycation end products (RAGE), which subsequently initiates various inflammatory and coagulation responses, thereby exacerbating atherosclerosis. The pathogenesis of atherosclerosis is also characterized by inflammation of the endothelial and perivascular tissues, encompassing processes such as vascular fibrosis and calcification [39]. Elevated concentrations of S100A8/A9 have been documented in patients with T2DM who have coronary artery disease, correlating with the severity of coronary artery disease [29]. Thus, increased levels of calprotectin in T2DM patients with macrovascular disease may indicate heightened vascular inflammation and enhanced plaque vulnerability. However, given the limited number of studies and the partial instability observed in sensitivity analyses, these findings warrant cautious interpretation.

In the present meta-analysis, circulating calprotectin levels did not exhibit a statistically significant difference between patients with T2DM presenting with microalbuminuria and those with normoalbuminuria. However, individual studies have reported heterogeneous findings. Schmaderer et al. [31] demonstrated significantly elevated serum MRP8/14 (calprotectin) levels in T2DM patients with microalbuminuria compared to normoalbuminuric individuals, positing that enhanced macrophage transmigration—evidenced by increased circulating MRP8/14—may signify an early event in the pathogenesis of diabetic microvascular injury. The systemic elevation of MRP8/14, a macrophage-derived inflammatory mediator, supports the hypothesis of chronic low-grade inflammation during the initial stages of diabetic nephropathy. Conversely, Gu et al. [23] observed no significant difference in calprotectin concentrations between T2DM patients with normoalbuminuria and those with microalbuminuria, although markedly higher levels were noted in individuals with macroalbuminuria. This stepwise increase across stages of albuminuria suggests that circulating calprotectin may serve as a marker of disease severity rather than an indicator of early renal involvement, indicating a potentially stronger correlation with advanced microcirculatory damage in diabetic kidney disease (DKD). Collectively, these discrepant findings underscore the dynamic role of MRP8/14 in the progression of DKD and may partially elucidate the lack of significant association observed in pooled analyses at the microalbuminuria stage. Evidence regarding urinary calprotectin in T2DM remains limited and inconsistent. The two available studies exhibited considerable heterogeneity and did not reveal a statistically significant pooled difference between T2DM patients and controls.

The strengths of this meta-analysis include adherence to the PRISMA guidelines, prospective registration with PROSPERO, a comprehensive literature search, a rigorous quality assessment, and extensive sensitivity and subgroup analyses. However, several limitations must be acknowledged. A majority of the included studies employed cross-sectional designs, which precludes the establishment of causal inferences. Notably, significant heterogeneity and evidence of small-study effects were observed, particularly within complication-specific analyses. Furthermore, incomplete adjustment for confounding variables, such as obesity, glycemic control, and medication usage, may have influenced the results. Lastly, the limited number of studies investigating urinary calprotectin and specific complications constrains the generalizability of these findings.

Conclusions

In conclusion, this systematic review and meta-analysis elucidates that circulating calprotectin levels are significantly elevated in individuals with Type 2 Diabetes Mellitus (T2DM), thereby reinforcing its potential role as a biomarker of chronic low-grade inflammation in diabetes. Preliminary evidence indicates that elevated calprotectin levels may be associated with macrovascular complications and peripheral neuropathy; however, these associations necessitate validation through larger, rigorously designed prospective studies. Future research should prioritize longitudinal assessments of its prognostic value and investigate the potential of calprotectin as a therapeutic target in the context of metabolic inflammation.

Acknowledgements
None.

Funding
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Conflict of interests
The authors declare that they have no conflict of interest regarding the publication of this article.

Ethical approval
This study is a systematic review and meta-analysis. No ethical approval is required.

Credit Author Statements
Roshan Kumar Mahat: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration.
Mritunjay Kumar Mishra: Methodology, Investigation, Data Curation, Validation, Writing - Review & Editing.
Nand Kishor Prasad Sah: Methodology, Writing - Review & Editing.
Sushil Yadav: Methodology, Writing - Review & Editing.
All authors have read and approved the final manuscript

References

1. Magliano DJ, Boyko EJ, committee IDFDAtes. IDF Diabetes Atlas (10th ed.). IDF diabetes atlas. Brussels: International Diabetes Federation; 2021.
2. Harding JL, Pavkov ME, Magliano DJ, Shaw JE, Gregg EW. Global trends in diabetes complications: a review of current evidence. Diabetologia. 2019;62(1):3-16. https://doi.org/10.1007/s00125-018-4711-2
3. Ehrchen JM, Sunderkötter C, Foell D, Vogl T, Roth J. The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. J Leukoc Biol. 2009;86(3):557-566. https://doi.org/10.1189/jlb.1008647
4. Altwegg LA, Neidhart M, Hersberger M, Müller S, Eberli FR, Corti R, et al. Myeloid-related protein 8/14 complex is released by monocytes and granulocytes at the site of coronary occlusion: a novel, early, and sensitive marker of acute coronary syndromes. Eur Heart J. 2007;28(8):941-948. https://doi.org/10.1093/eurheartj/ehm078
5. Healy AM, Pickard MD, Pradhan AD, Wang Y, Chen Z, Croce K, et al. Platelet expression profiling and clinical validation of myeloid-related protein-14 as a novel determinant of cardiovascular events. Circulation. 2006;113(19):2278-2284. https://doi.org/10.1161/CIRCULATIONAHA.105.607333
6. Jarlborg M, Courvoisier DS, Lamacchia C, Martinez Prat L, Mahler M, Bentow C, et al. Serum calprotectin: a promising biomarker in rheumatoid arthritis and axial spondyloarthritis. Arthritis Res Ther. 2020;22(1):105. https://doi.org/10.1186/s13075-020-02190-3
7. Gran S, Honold L, Fehler O, Zenker S, Eligehausen S, Kuhlmann MT et al. Imaging, myeloid precursor immortalization, and genome editing for defining mechanisms of leukocyte recruitment in vivo. Theranostics. 2018;8(9):2407-2423. https://doi.org/10.7150/thno.23632
8. Hessian PA, Edgeworth J, Hogg N. MRP-8 and MRP-14, two abundant Ca(2+)-binding proteins of neutrophils and monocytes. J Leukoc Biol. 1993;53(2):197-204.
9. Sorg C. The calcium binding proteins MRP8 and MRP14 in acute and chronic inflammation. Behring Inst Mitt. 1992;(91):126-137.
10. Vogl T, Tenbrock K, Ludwig S, Leukert N, Ehrhardt C, van Zoelen MA, et al. Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. Nat Med. 2007;13(9):1042-1049. https://doi.org/10.1038/nm1638
11. Ghavami S, Rashedi I, Dattilo BM, Eshraghi M, Chazin WJ, Hashemi M, et al. S100A8/A9 at low concentration promotes tumor cell growth via RAGE ligation and MAP kinase-dependent pathway. J Leukoc Biol. 2008;83(6):1484-1492. https://doi.org/10.1189/jlb.0607397
12. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. J Clin Epidemiol. 2021;134:178-189. https://doi.org/10.1016/j.jclinepi.2021.03.001
13. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (accessed on: 05/12/2025)
14. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: Systematic reviews of etiology and risk . In: Aromataris E, Munn Z (Editors). Joanna Briggs Institute Reviewer's Manual. The Joanna Briggs Institute, 2017.
15. Abbas A, Hefnawy MT, Negida A. Meta-analysis accelerator: a comprehensive tool for statistical data conversion in systematic reviews with meta-analysis. BMC Med Res Methodol. 2024;24(1):243. https://doi.org/10.1186/s12874-024-02356-6
16. Pedersen L, Nybo M, Poulsen MK, Henriksen JE, Dahl J, Rasmussen LM. Plasma calprotectin and its association with cardiovascular disease manifestations, obesity and the metabolic syndrome in type 2 diabetes mellitus patients. BMC Cardiovasc Disord. 2014;14:196. https://doi.org/10.1186/1471-2261-14-196
17. Furuya-Kanamori L, Barendregt JJ, Doi SA. A new improved graphical and quantitative method for detecting bias in meta-analysis. Int J Evid Based Healthc. 2018;16(4):195-203. https://doi.org/10.1097/

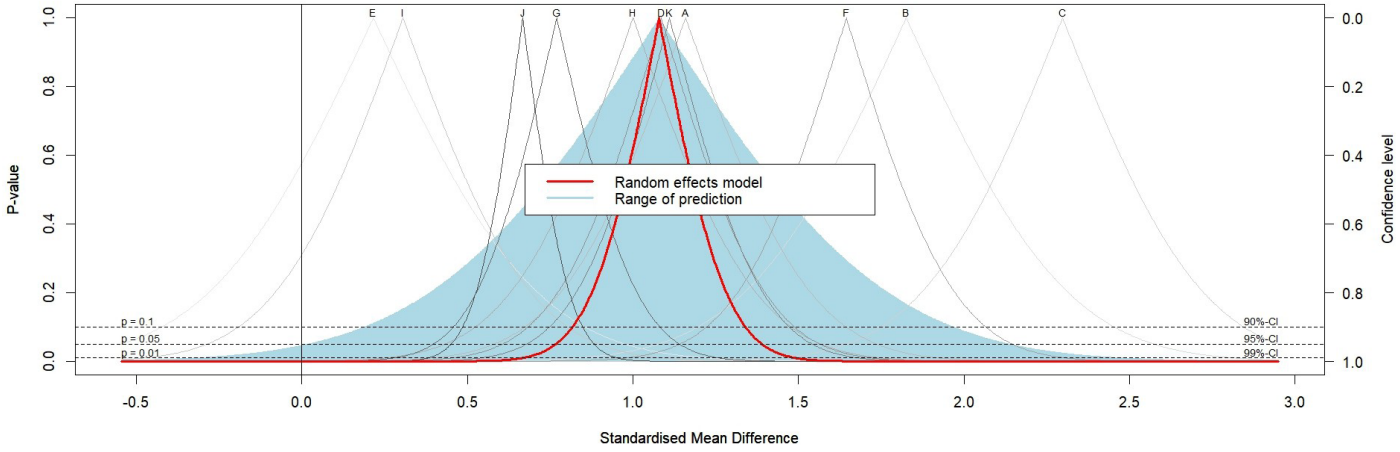
XEB.0000000000000141
18. Afify MA, El Khawaga SSH, Adly NN, Awadein MA, El Din AMB. The relationship between serum calprotectin and peripheral neuropathy in a sample of Egyptian type 2 diabetic patients. Ain Shams Med J. 2022;73(1):153-169.
19. El-Adawy NM, El-Hini SH, Hassan EM, Abd El Wahab EA. Calprotectin as biomarker in diabetic patients with peripheral neuropathy. MJMR. 2017;28(3):25-27.
20. El-Hafez FFA, Nsr-Allah AAM, Mohamed AKAE, Ahmed AM, Mahmoud AA. Novel biomarker serum calprotectin for early diagnosis of diabetic peripheral neuropathy in type 2 diabetes patients. Egypt J Hosp Med. 2021;82(2):379-385.
21. Gao H, Xu J, He L, Meng H, Hou J. Calprotectin levels in gingival crevicular fluid and serum of patients with chronic periodontitis and type 2 diabetes mellitus before and after initial periodontal therapy. J Periodontal Res. 2021;56(1):121-130. https://doi.org/10.1111/jre.12800
22. Gobejishvili L, Vatsalya V, Avila DV, Feygin YB, McClain CJ, Mokshagundam S, et al. Association of Circulating Markers of Microbial Translocation and Hepatic Inflammation with Liver Injury in Patients with Type 2 Diabetes. Biomedicines. 2024;12(6):1227. https://doi.org/10.3390/biomedicines12061227
23. Gu W, Cui J, Shi M, Zhang R, Tan D, Zhang X, et al. Serum myeloid-related protein 8/14 is a potential predictor of diabetic kidney disease. Discov Med. 2020;30(160):97-105.
24. Halawa M, Abushady MM, Abdelsalam Besibes MM, Basabbea AS, Mohamed NR. Plasma and urinary calprotectin levels in type 2 diabetic patients with peripheral arterial disease. Egypt J Hosp Med. 2020;81(6):2222-2226.
25. Hamzah MI, Abdul Kareem IA. Evaluation of calprotectin and catalase enzyme in type 2 diabetes mellitus. Biochem Cell Arch. 2020;20(2):6391–6395.
26. Hassan HN, Shamran SG, Albadry MAZ, Al-Fahham AA. Evaluation of serum levels of calprotectin, lactoferrin and zinc in patients with type II diabetes mellitus. Wiad Lek. 2025;78(2):288–294. https://doi.org/10.36740/WLek/201198
27. Lylloff L, Bathum L, Madsbad S, Grundtvig JLG, Nordgaard-Lassen I, Fenger M. S100A8/A9 (calprotectin), interleukin-6, and C-reactive protein in obesity and diabetes before and after Roux-en-Y gastric bypass surgery. Obes Facts. 2017;10(5):386–395. https://doi.org/10.1159/000478097
28. Murad HAS, Ghabrah TMA, Bakarman MAM, Rafeeq MM, Eid SS. Is urinary calprotectin a useful tool to detect the therapeutic control of type II diabetes? Int J Pharmacol. 2019;15(2):295–300. https://doi.org/10.3923/ijp.2019.295.300
29. Peng WH, Jian WX, Li HL, Hou L, Wei YD, Li WM, et al. Increased serum myeloid-related protein 8/14 level is associated with atherosclerosis in type 2 diabetic patients. Cardiovasc Diabetol. 2011;10:41. https://doi.org/10.1186/1475-2840-10-41
30. Salem IA, Abdulsattar SA, Alrubaye HF. Serum calprotectin level in type 2 diabetic patients with and without diabetic peripheral neuropathy: A comparison study. Mustansiriya Med J. 2024;23(2):55–60. https://doi.org/10.4103/mj.mj_11_24
31. Schmaderer C, Kemmner S, Burkhardt K, Heemann U, Baumann M. Serum myeloid-related protein 8/14 complex is associated with microalbuminuria in patients with type 2 diabetes. Ther Adv Cardiovasc Dis. 2014;8(3):80-88. https://doi.org/10.1177/1753944714528270
32. Tabur S, Korkmaz H, Ozkaya M, Aksoy SN, Akarsu E. Is calprotectin a novel biomarker of neuroinflammation in diabetic periferal neuropathy? Diabetol Metab Syndr. 2015;7:36. https://doi.org/10.1186/s13098-015-0030-7
33. Velayutham R, Nair PP, Adole PS, Mehalingam V. Association of serum calprotectin with peripheral neuropathy in patients with type 2 diabetes mellitus. J Family Med Prim Care. 2021;10(4):1602-1606. https://doi.org/10.4103/jfmpc.jfmpc_1165_20
34. Wang Z, Ni X, Zhang L, Sun L, Zhu X, Zhou Q, et al. Toll-Like Receptor 4 and Inflammatory Micro-Environment of Pancreatic Islets in Type-2 Diabetes Mellitus: A Therapeutic Perspective. Diabetes Metab Syndr Obes. 2020;13:4261-4272. https://doi.org/10.2147/DMSO.S279104
35. Yan SF, Ramasamy R, Naka Y, Schmidt AM. Glycation, inflammation, and RAGE: a scaffold for the macrovascular complications of diabetes and beyond. Circ Res. 2003;93(12):1159-1169. https://doi.org/10.1161/01.RES.0000103862.26506.3D
36. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. Nat Rev Immunol. 2011;11(2):98-107. https://doi.org/10.1038/nri2925
37. Pickup JC. Inflammation and activated innate immunity in the pathogenesis of type 2 diabetes. Diabetes Care. 2004;27(3):813-823. https://doi.org/10.2337/diacare.27.3.813
38. Catalán V, Gómez-Ambrosi J, Rodríguez A, Ramírez B, Rotellar F, Valentí V, et al. Increased levels of calprotectin in obesity are related to macrophage content: impact on inflammation and effect of weight loss. Mol Med. 2011;17(11-12):1157-1167. https://doi.org/10.2119/molmed.2011.00144
39. Bourgonje AR, Bourgonje MF, Sokooti S, la Bastide-van Gemert S, Nilsen T, Hidden C, et al. Plasma Calprotectin and New-onset Type 2 Diabetes in the General Population: A Prospective Cohort Study. J Clin Endocrinol Metab. 2024;110(1):e150-e159. https://doi.org/10.1210/clinem/dgae130

Supplementary files

Supplementary Table 1: Search strategies in databases.

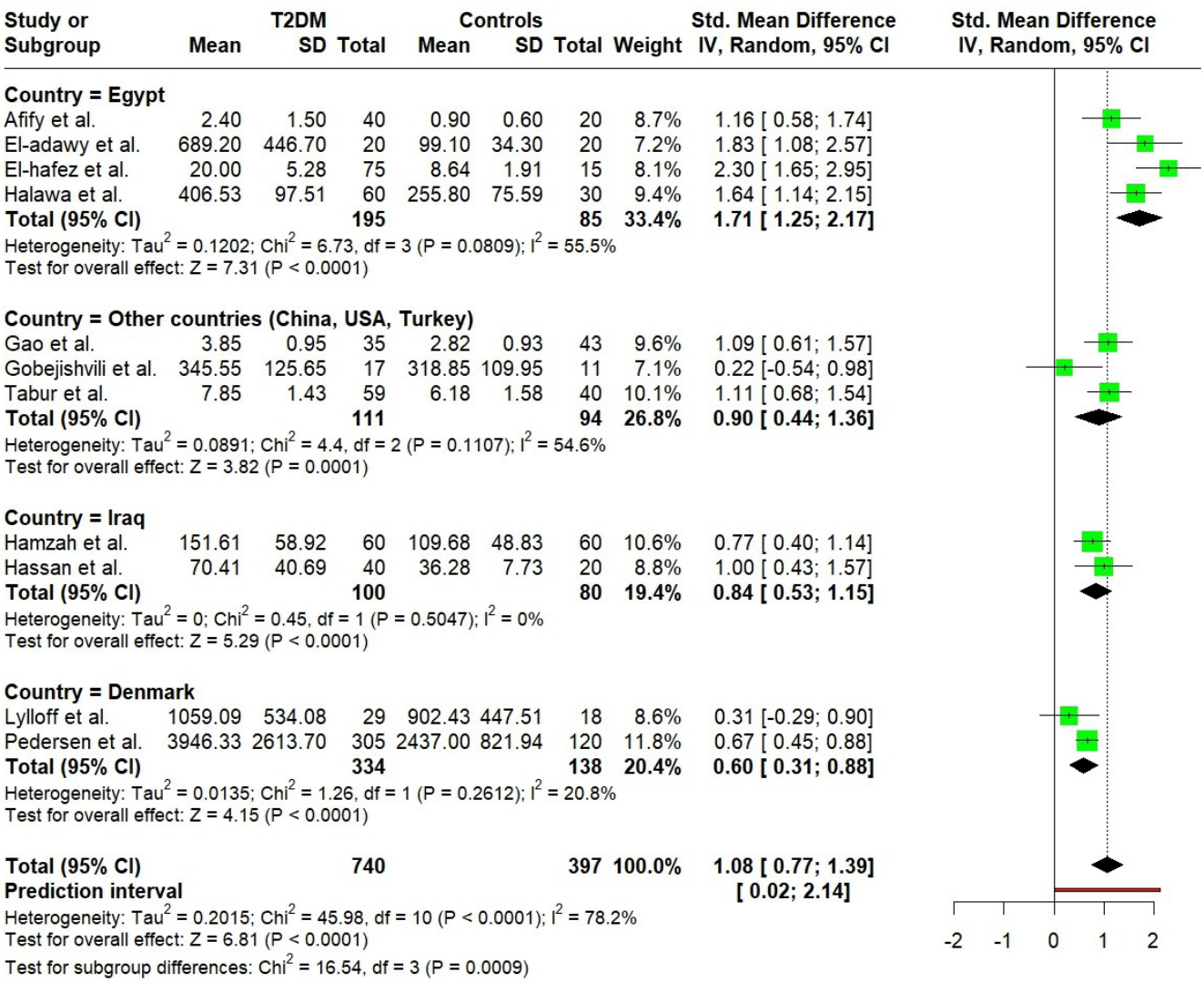
Databases	Search terms	No of articles identified
PubMed/Medline	("type 2 diabetes mellitus"[All Fields] OR "diabetes mellitus"[All Fields] OR "type 2 diabetes"[All Fields] OR "diabetes"[All Fields] OR "diabetics"[All Fields] OR "type 2 DM"[All Fields] OR "T2DM"[All Fields] OR "T2D"[All Fields]) AND ("calprotectin"[All Fields] OR "myeloid-related protein 8/14"[All Fields] OR "MRP 8/14"[All Fields] OR "S100A8/A9"[All Fields] OR "leukocyte protein L1"[All Fields])	171
Scopus	TITLE-ABS-KEY (("type 2 diabetes mellitus" OR "diabetes mellitus" OR "type 2 diabetes" OR "diabetes" OR "diabetics" OR "type 2 DM" OR "T2DM" OR T2D) AND ("calprotectin" OR "myeloid-related protein 8/14" OR "MRP 8/14" OR "S100A8/A9" OR "leukocyte protein L1"))	219
Web of Science	("type 2 diabetes mellitus" OR "diabetes mellitus" OR "type 2 diabetes" OR "diabetes" OR "diabetics" OR "type 2 DM" OR "T2DM" OR T2D) AND ("calprotectin" OR "myeloid-related protein 8/14" OR "MRP 8/14" OR "S100A8/A9" OR "leukocyte protein L1")	205

Supplementary Figure 1: Drapery plot depicting the meta-analysis results based on study-specific p-value functions in T2DM.

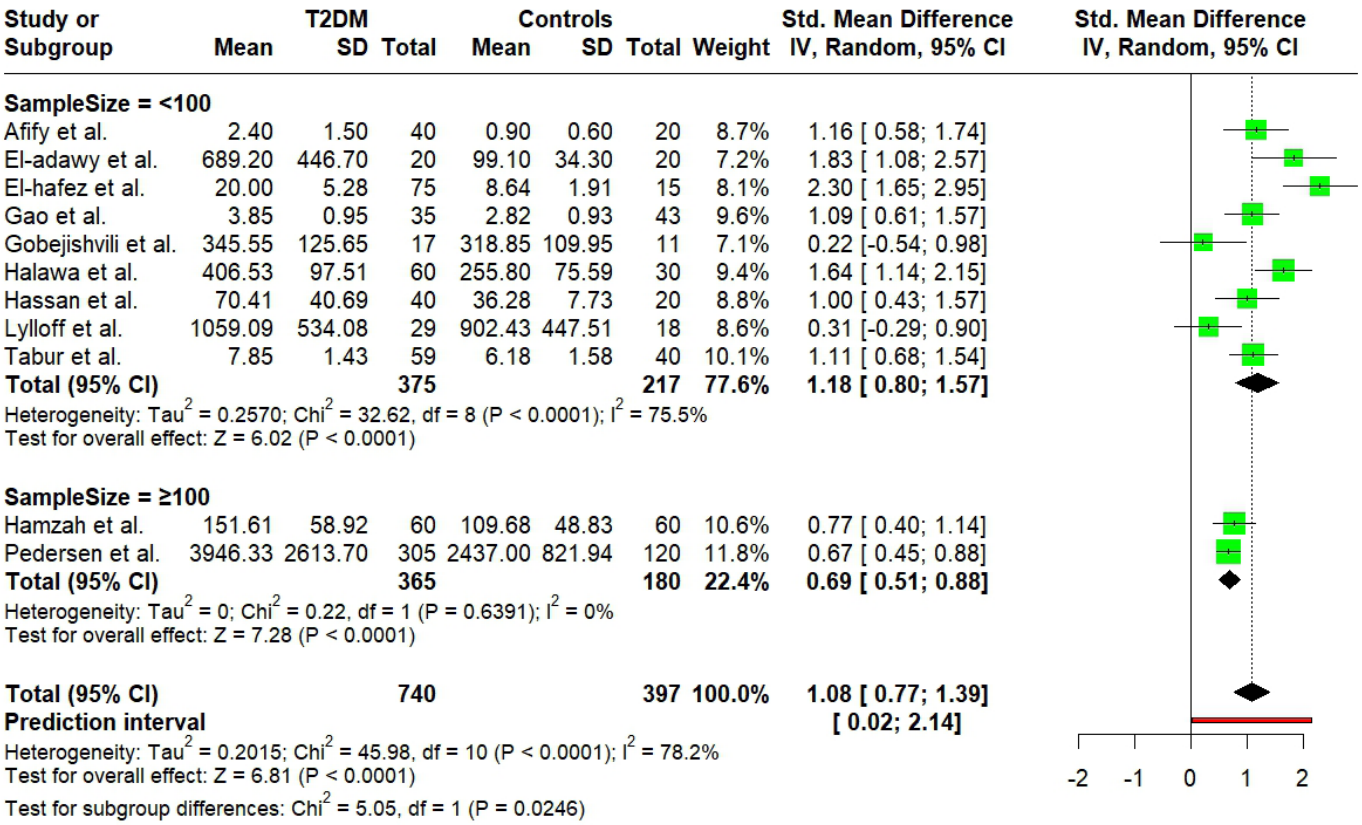


A: Afify et al. [18]; B: El-adawy et al. [19]; C: El-hafez et al. [20]; D: Gao et al. [21]; E: Gobejishvili et al. [22]; F: Halawa et al. [24]; G: Hamzah et al. [25] H: Hassan et al. [26]; I: Lylloff et al. [27]; J: Pedersen et al. [16]; K: Tabur et al. [32]

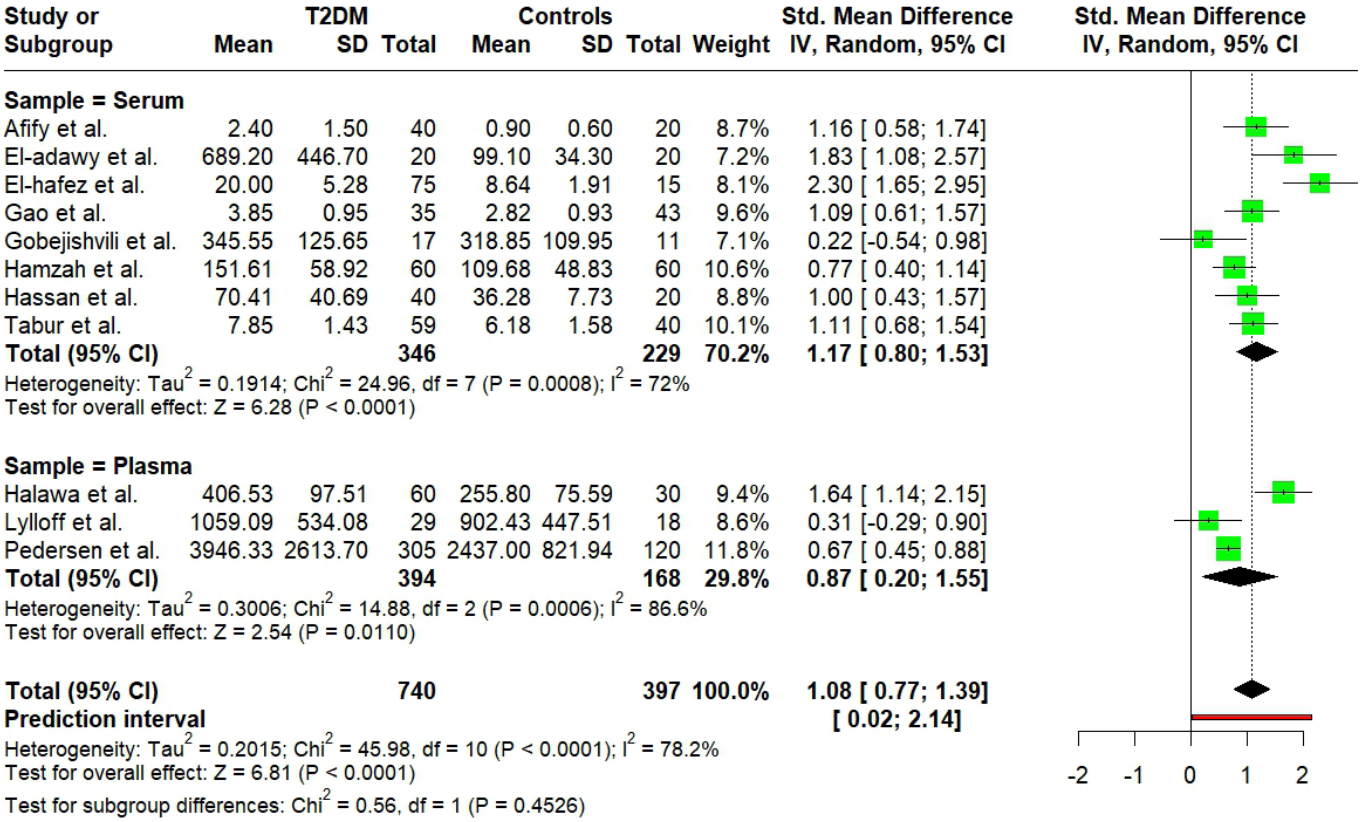
Supplementary Figure 2: Subgroup analysis by country.



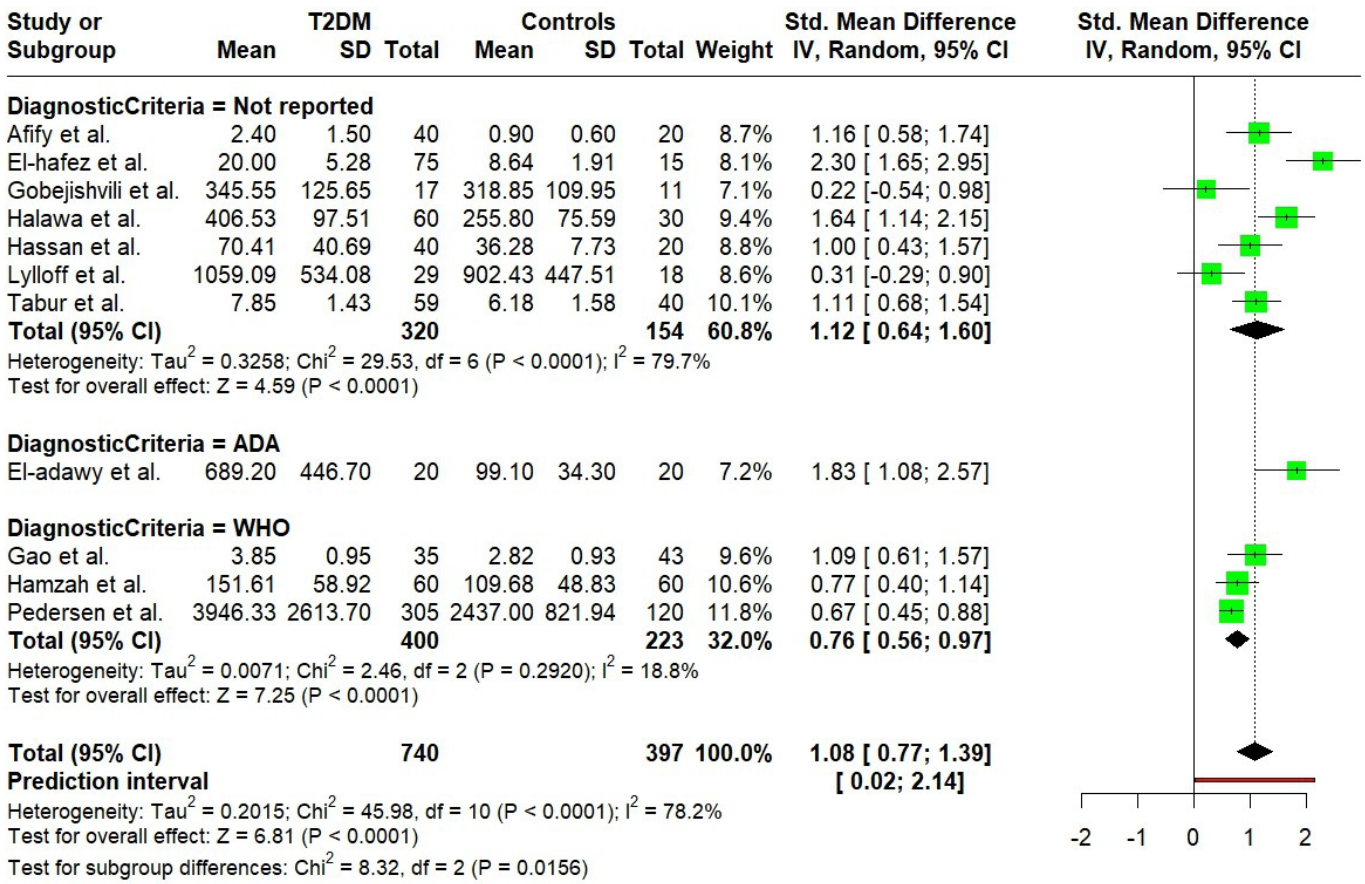
Supplementary Figure 3: Subgroup analysis by sample size.



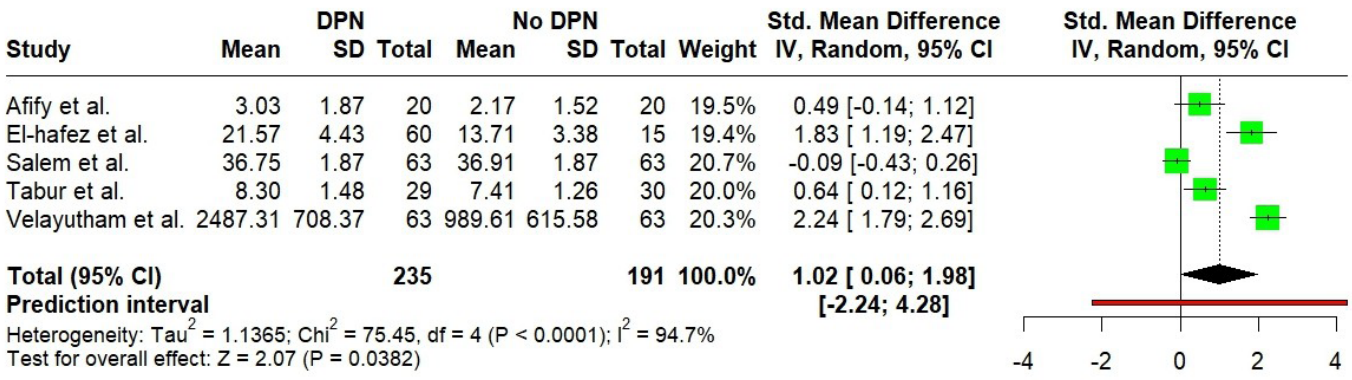
Supplementary Figure 4: Subgroup analysis by sample type.



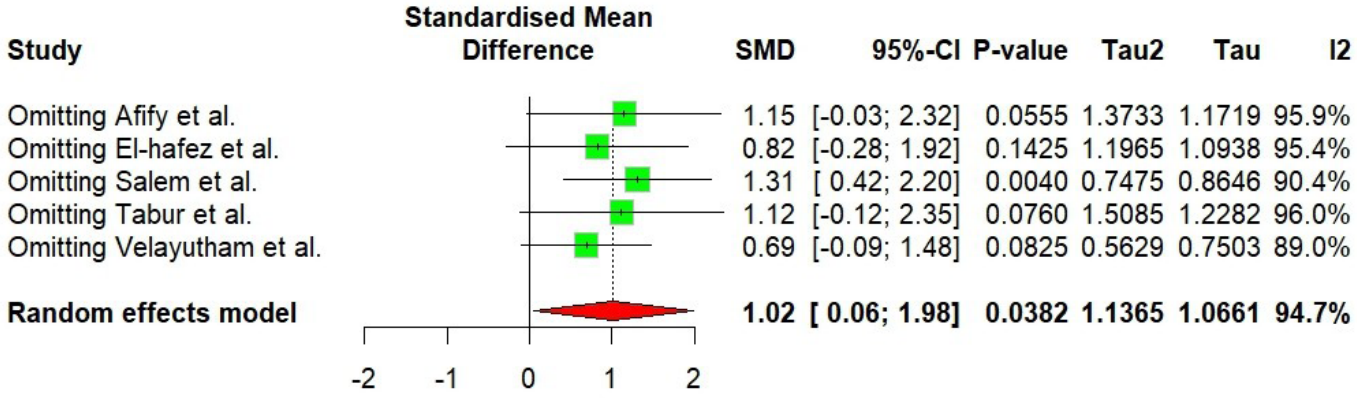
Supplementary Figure 5: Subgroup analysis by diagnostic criteria.



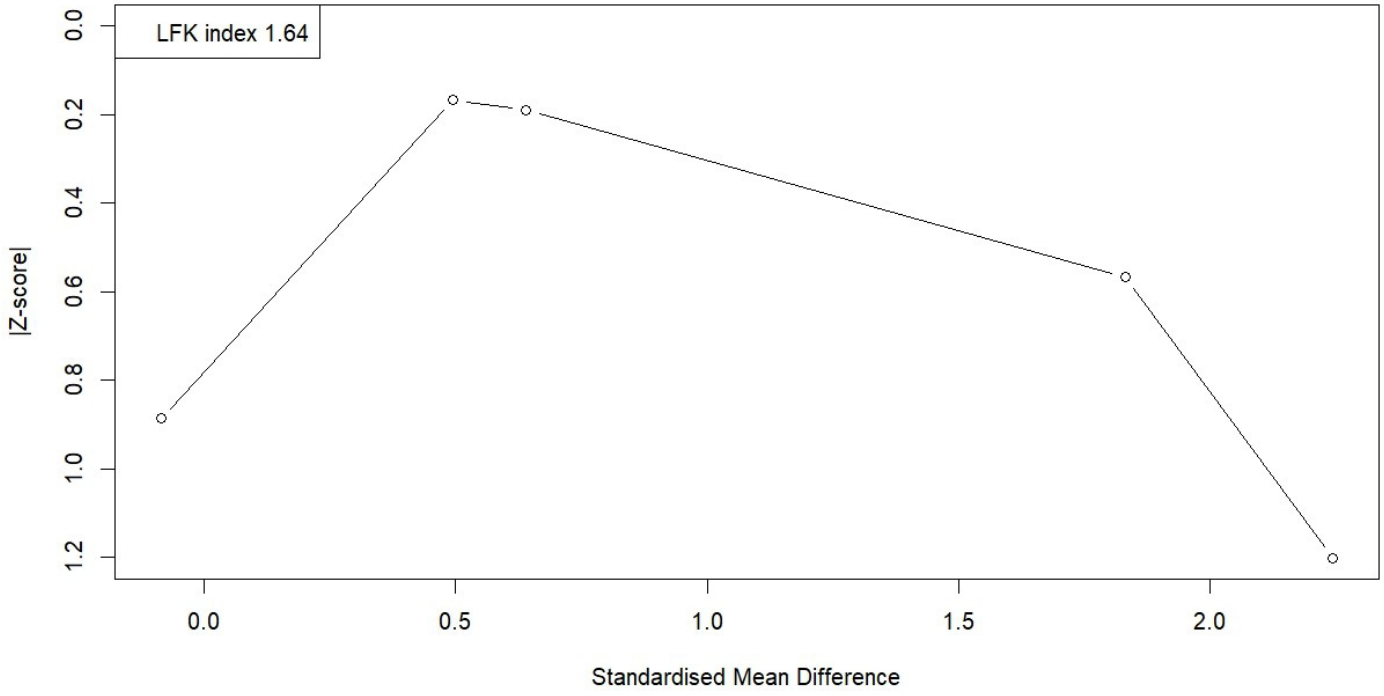
Supplementary Figure 6: Forest plot showing comparison of circulating calprotectin between individuals with and without diabetic peripheral neuropathy. DPN: diabetic peripheral neuropathy.



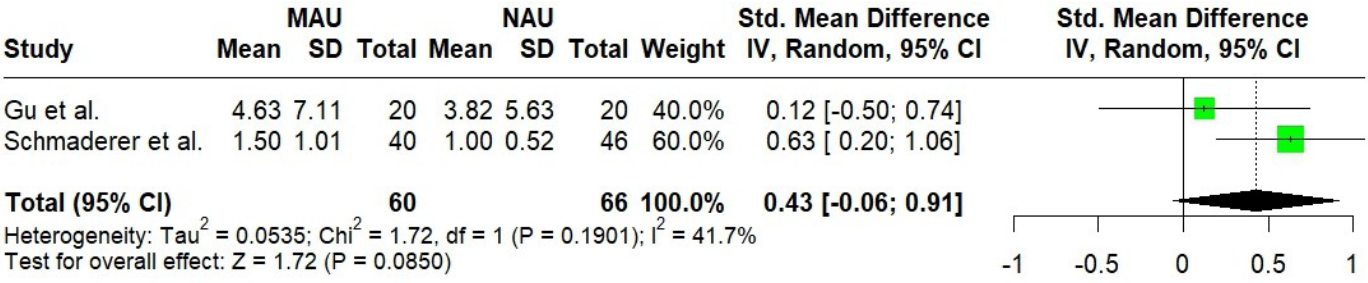
Supplementary Figure 7: Leave-one-out sensitivity analysis illustrating the influence of individual studies on the pooled estimate for diabetic peripheral neuropathy.



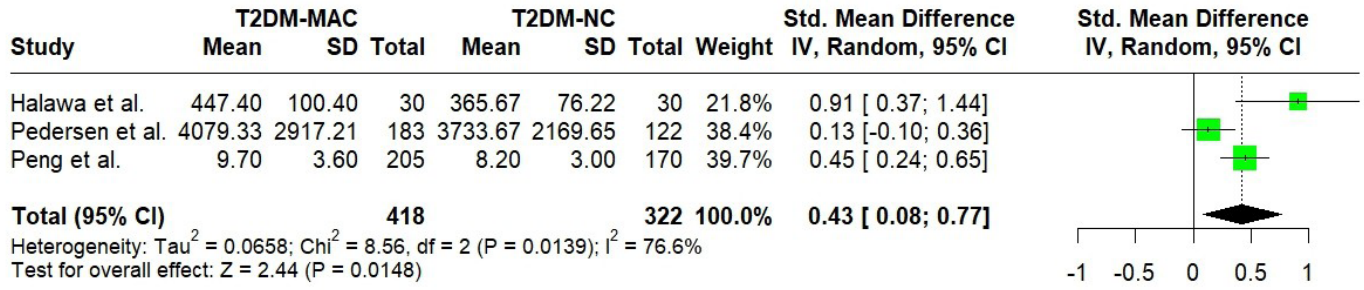
Supplementary Figure 8: Doi plot illustrating potential publication bias in diabetic peripheral neuropathy.



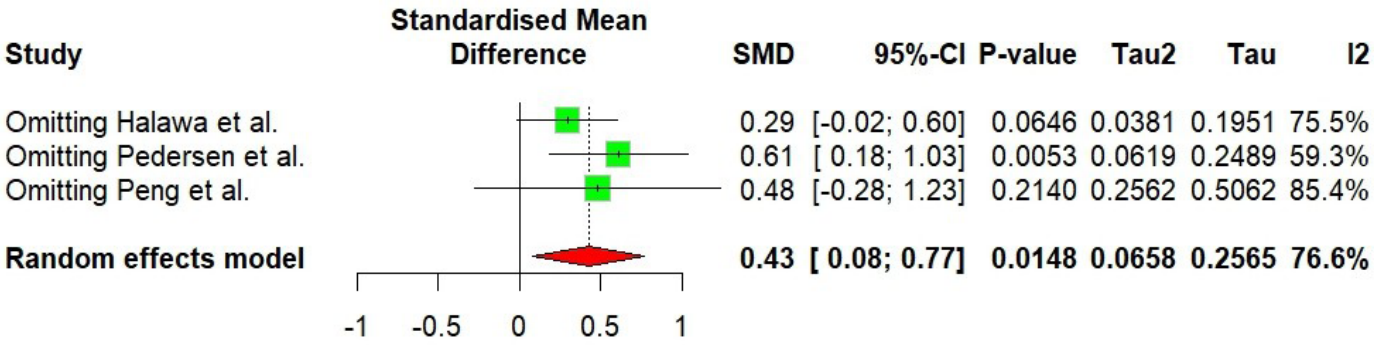
Supplementary Figure 9: Forest plot illustrating differences in circulating calprotectin levels between T2DM patients with normoalbuminuria (NAU) and microalbuminuria (MAU).



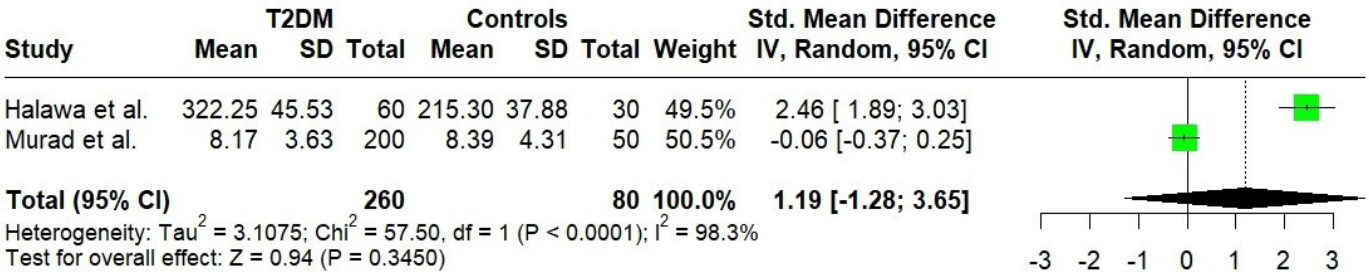
Supplementary Figure 10: Forest plot comparing circulating calprotectin levels between T2DM patients with macrovascular complications and those without macrovascular complications. T2DM-MAC: Type 2 diabetes mellitus with macrovascular complications; T2DM-NC: Type 2 diabetes mellitus without complications.



Supplementary Figure 11: Leave-one-out sensitivity analysis illustrating the influence of individual studies on the pooled estimate for T2DM with macrovascular complications.



Supplementary Figure 12: Forest plot showing comparison of urinary calprotectin between T2DM and controls.



Review article

Development of a Robust Point of Care Testing Program in Newfoundland and Labrador Using Lean Method of Rapid Process Improvement Workshop (RPIW)

Vinita Thakur[1,2*], Olatunji Anthony Akerele[1]

[1]*Department of Laboratory Medicine, Health Sciences Center, Newfoundland and Labrador Health Services, St. John's, Newfoundland and Labrador, Canada*

[2]*Faculty of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada*

Article Info

[*]***Corresponding Author:***
Vinita Thakur, M.Sc, Ph.D, DipRCPATH, NRCC, FADLM, Black Belt Lean Six Sigma,
Provincial Division Head and Scientific Director, Point of Care Testing Program
& Clinical Assistant Professor, Pathology and Laboratory Medicine Program,
Faculty of Medicine, Memorial University, Health Sciences Centre
Office 1J442, 300 Prince Philip Dr., St. John's, NL, CAN, A1B 3V6
Telephone: 709-777-6375 c:709-682-6326
E-mail: Vinita.Thakur@nlhealthservices.ca
Website: NLHealthServices.ca

Keywords

RPIW, Healthcare, Lean, Six Sigma and Point-of-care test

Abstract

The implementation of a provincial Point-of-Care Testing (POCT) program across Newfoundland and Labrador (NL), Canada, is designed to standardize verification processes and ensure consistent analytical performance of POCT analysers across diverse healthcare settings. The program is developed using a "Test and Treat" model to support both remote rural communities and urban populations. Many provincial locations operate as standalone POCT sites hence vendors are required to provide site-specific support services. Selection of analytically robust and cost-effective POCT devices is prioritized to support sustainability and quality patient care. Extensive project planning include Lean Six Sigma methodologies, including Rapid Process Improvement Workshop (RPIW), to standardize workflows and improve operational efficiency. Carefully selected Six Sigma team members participate in collaborative brainstorming sessions to develop streamlined verification and installation strategies across 23 provincial POCT sites. Each analyzer is assigned a Provincial Equipment Unit Identification Number (PEUIN) to enhance traceability and inventory management. Primary verification is conducted centrally at St. Clare's Mercy Hospital, St. John's, while secondary verification is completed at individual installation sites to ensure site-specific performance validation. Harmonized installation of identical analysers support a standardized and quality-driven approach to provincial POCT expansion. Structured utilization management processes are not yet fully established within NL Health Services. Development of utilization monitoring tools to evaluate financial implications and cost-effectiveness is ongoing. Quality improvement initiatives are continuously monitored using key performance indicators. Additionally, the implementation of the EPIC hospital information system in April 2026 provides opportunities for further standardization and integration of POCT processes across the province.

Introduction

Point-of-Care Testing (POCT) refers to diagnostic testing conducted near the patient's bedside, enabling faster turnaround times compared to traditional laboratory services [1]. Traditional laboratory testing follows a multi-step process that begins with sample collection at the bedside or clinic, followed by transportation to a centralized laboratory, often located at a significant distance. Once there, samples undergo multiple stages of processing. This extended timeline can delay diagnosis and treatment, ultimately hindering timely clinical decision-making. By delivering rapid results at the site of care, POCT supports immediate clinical decision making and timely patient management, reducing delays and improving outcomes and overall satisfaction with healthcare services [2]. The practical challenges of managing POCT compared to centralized laboratory testing, and its impact on improving healthcare outcomes have been debated [3]. However, POCT has demonstrated clear clinical benefits of facilitating expedited diagnoses, faster treatment initiation, and enhanced patient care in acute care settings [4].

The COVID-19 pandemic underscored the critical value of having a robust POCT infrastructure in the healthcare system [5]. The global health crisis highlighted the urgent need for rapid, decentralized testing capabilities, particularly in settings with limited access to centralized laboratories. The experience of the pandemic validated the importance of well-equipped POCT programs with modern instrumentation and appropriately trained personnel to help mitigate community transmission. As such, many diagnostics companies continue to invest significantly in developing advanced POCT technologies.

The province of Newfoundland and Labrador (NL) has one healthcare authority, Newfoundland and Labrador Health Services (NLHS), tasked with the administration & delivery of healthcare services across the province. NLHS is divided into 5 zones to balance centralized province-wide efficient healthcare delivery. These zones are named Eastern (Urban, and Rural), Central, Western, and Labrador & Grenfell zones. Laboratory services within NLHS are provided across the province through major hospitals and regional health centers.

NL is characterized by its rugged North Atlantic coastline and harsh winter conditions. NL is home to many remote and rural communities, often located 200 to 300 kilometers away from major hospitals. This realization led the provincial government to recognize the urgent need for a strong POCT network in the province. In response to these geographical and demographic challenges, and in alignment with provincial directives, Department of Pathology and Laboratory Medicine initiated a strategic transition from a centralized laboratory model to a province-wide POCT-based service delivery model. This approach aims to provide timely and uninterrupted diagnostic services to even the most remote communities across NL. More so, the provision of a timely, robust, and standardized province-wide medical care for the aging and widely dispersed population aligns with the quality improvement initiatives across all healthcare zones within NLHS. Central to building this system was the implementation of a cost-effective lean process, designed to streamline operations and enhance efficiency.

As the demand for POCT continued to evolve in the province, the division of Pathology and Laboratory Medicine (PLM) services initiated the process of phasing out routine laboratory testing from smaller rural facilities and replacing it with a POCT-centric model. This transition includes regionalizing routine testing and centralizing specialized diagnostics services at hub sites. Under this framework, outpatient samples such as blood and urine are transported to central laboratories for routine analysis, while POCT is used to rapidly deliver test results for immediate intervention directly at the point of care. To further support this initiative, the POCT team members undertook the responsibility of developing unified policies, procedures, and workflows. NLHS POCT covered mainly 23 rural sites in addition to other isolated coastal community sites. Each project for installation of various POCT analysers across 23 provincial rural sites constituted five phases of DMADV (Define, Measure, Analyze, Design and Verify) framework guided by Lean Six Sigma principles.

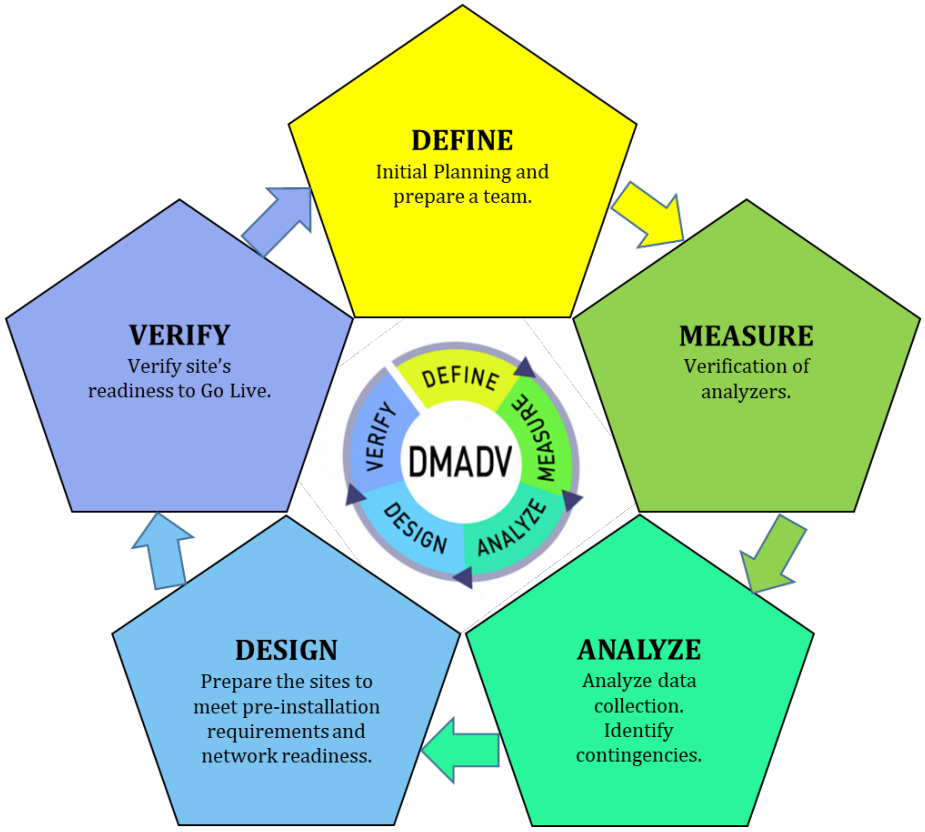
The Lean Six Sigma is a quality improvement methodology designed to systematically reduce defects and minimize variation in processes. Originally developed for the automotive and manufacturing industries, it's application in healthcare is relatively recent but increasingly impactful [7]. Today, Lean Six Sigma has established a strong presence in healthcare systems, driving operational excellence across various domains. Recent literature underscores its effectiveness in addressing operational challenges in emergency departments, cardiology, surgery, intensive care units, pediatrics, pharmacy, radiology, and laboratory services [6 – 11]. In addition to the standard DMADV framework, we employed another Lean Six Sigma tool; the Rapid Process Improvement Workshop (RPIW), to establish a strong foundation for our POCT program [12].

Description of sigma tools

Pentagon Model of DMADV

Each project plan was broadly based on pentagon model, DMADV, a Lean Six Sigma planning strategy for smooth execution of project activities. Each phase describes the activities involved as shown in Figure 1 below. "Pentagon model" term has been used by us for the first time due to the design of this framework.

Figure 1: A Pentagon model of DMADV process of NL-POCT program.



Description of Rapid Process Improvement Workshop (RPIW) Sigma tool

A RPIW is a Lean Six Sigma tool designed to assess and improve utilization management and existing workflows before launching a new program. It begins with the collection of baseline data on current practices, followed by a focused five-day workshop involving key stakeholders. During the workshop, participants collaboratively analyze current processes, identify inefficiencies, and develop a future-state model along with a detailed action plan that includes measurable outcomes and evaluation strategies [12-14]. Key activities in this framework include pre-workshop planning, data collection, project charter development, volunteer or team member selection, defect identification, and iterative planning to remove process inefficiencies (Muda) [14]. Typically, a 6–10 weeks pre-work phase precedes the workshop to define waste and establish improvement goals. On Day 1, participants receive RPIW training and review collected data. Day 2 involves process mapping (current and future state), gap analysis, and identification of challenges. The remaining days are used for action planning, implementation trials, and evaluation, guided by the Plan-Do-Study-Act (PDSA) cycle [15]. Follow-up assessments are conducted at 30, 60, and 90 days to monitor and sustain progress. The U.S. Veterans Health Administration (VHA) used RPIW to implement its eScreening program, a mobile, automated health screening platform for veterans [12].

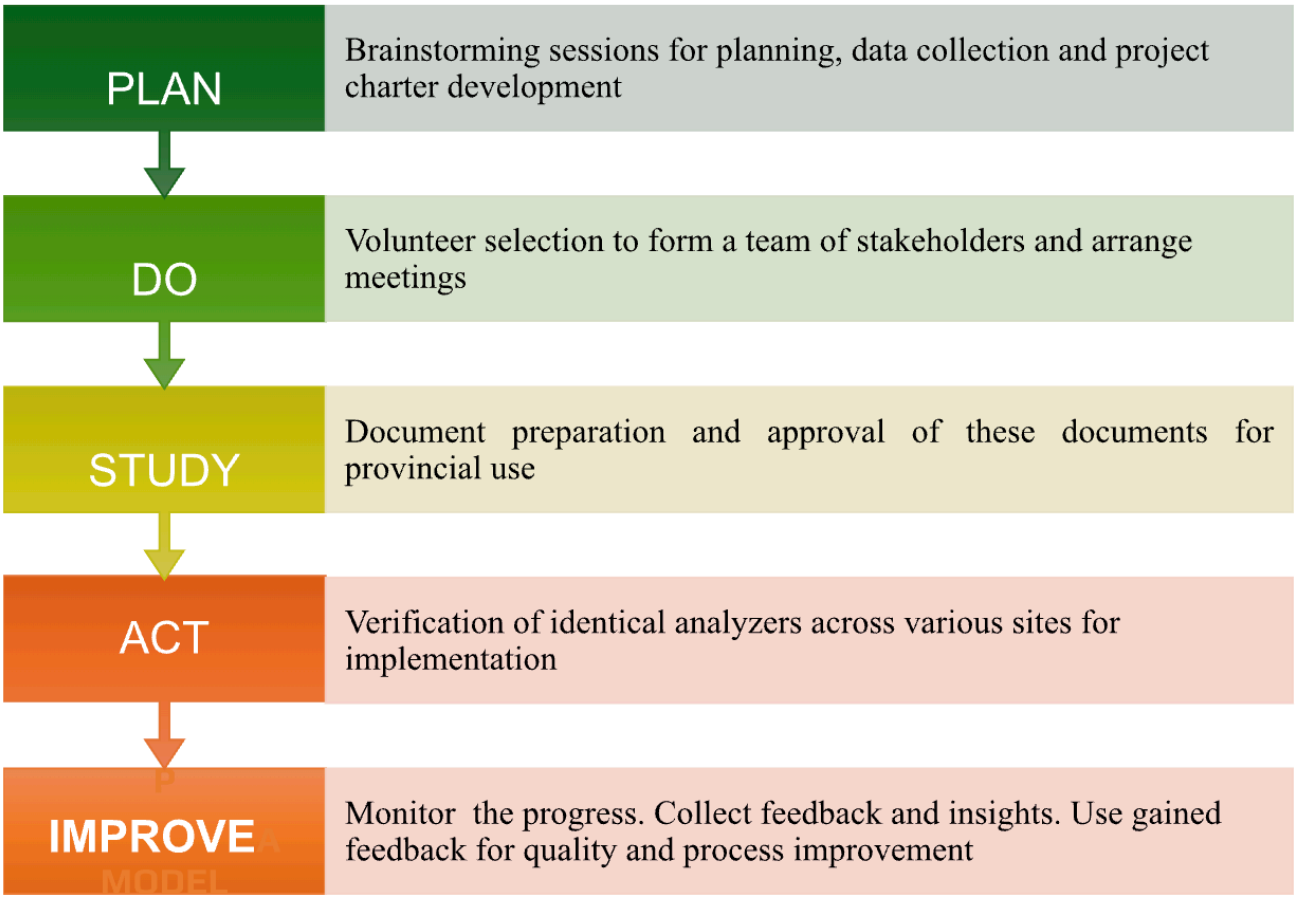
Development of RPIW framework for NLHS-POCT

A typical RPIW plan was framed for developing our process. The Eastern Urban zone POCT center at St Clare's Mercy Hospital (SCH) served as the primary referral center supporting multiple POCT sites (n=23) across the province. Project initiation and implementation within the division closely followed the RPIW framework to ensure structured and effective execution. Project charter was developed and signed. The POCT team was constituted which includes committed members who volunteered for various project assignments, contributing to collaborative progress. This initiative was strongly supported by the Medical Scientific leadership, including the Division Head of POCT and the Clinical Chief of the Pathology and Laboratory Medicine Program, who actively oversaw and guided project milestones. Regular team meetings were held to develop and refine strategies aimed at implementing a streamlined and standardized verification process across all provincial POCT sites.

PDSAI structure to portray RPIW

The utilization improvement plan, a typical PDSAI (Plan, Do, Study, Act and Improve) framework, was structured into five key steps as illustrated in Figure 2. We added an "Improve" step in typical PDSA framework to make this model more efficient for monitoring quality improvement.

Figure 2: A typical PDSAI model to mirror the RPIW plan for developing NLHS-POCT program.



A detailed description of PDSAI framework is given below.

1. **Initial Brainstorming Sessions (P):** Collaborative meetings were held with provincial POCT staff, the IT team, quality assurance personnel, and other relevant stakeholders to establish a unified direction. Project charter was developed and signed.
2. **Requirement Identification and Task Delegation (D):** The team identified the verification requirements for each POCT instrument and outlined the steps necessary for deployment. Specific responsibilities were assigned to team members, each with a defined timeline for completion.
3. **Document Preparation and Approval (S):** Essential documentation was prepared, including the new equipment accession form, equipment transfer policy, pre-move checklist, master inventory list, verification protocols, and standard operating procedures (SOPs). These documents were studied, reviewed and approved by designated signatories prior to implementation.
4. **Verification Phases (A):** Primary verification was conducted by the POCT team at central sites such as St. Clare's Mercy Hospital (SCH) or the Health Sciences Centre (HSC) St. John's. Secondary verification was carried out at the final site of installation and were compared with site specific analysers before the device's approval for clinical use.
5. **Continuous Improvement (I):** Feedback and insights gained for the implementation process were used to generate ideas for quality and process improvement. These adaptations ensured the model remains responsive and effective throughout the project lifecycle.

Project Model of NLHS-POCT Program

A project charter was developed to guide all provincial POCT analyzer installations, accompanied by a Gantt chart outlining the overall process plan and implementation timeline. Regular scheduled meetings were conducted to monitor progress and ensure timely execution of each project phase. Analysers were deployed to designated sites either in acute care and community hospitals or at remote standalone POCT sites. These instruments were introduced either as replacements for outdated equipment as the backup for main Chemistry analysers or as new additions to expand the POCT services. In adherence to Good Laboratory Practices (GLP), all necessary documentation was meticulously completed for each installation. This included the master inventory list, equipment accession forms,

equipment transfer policy, pre-move checklist, SOPs and quick reference guide. Each analyzer was assigned a unique PEUIN to facilitate standardized tracking and management.
Initially, installations of analysers took place within laboratory settings and were operated by laboratory technologists until the program matured into a fully integrated point-of-care model.

Provincial Equipment Unique Identification Number (PEUIN)
Each newly installed analyzer was assigned a unique PEUIN in addition to the existing inventory asset number generated by the Biomedical Engineering Department of each health zone. The PEUIN was developed specifically for the standardized identification, verification, and inventory management of POCT equipment across the province. This helped in timely identification of sites and type of equipment if any failure or interruption in services occurred.
The PEUIN was systematically structured using the following format:

- The first three letters of the site name (e.g., SCH for St. Clare's Mercy Hospital)
- The division code (POC for Point of Care)
- The analyzer trade name or model (e.g., ABL)
- A sequential number (e.g., 1, 2, 3…)
- The two-letter code for the device type (e.g., BG for Blood Gas Analyzer)

For example: SCH-POC-ABL1BG represents the first ABL blood gas analyzer installed at St. Clare's Mercy Hospital within the POCT division. This unique identifier (PEUIN) was incorporated into the verification protocol for each instrument to ensure proper alignment between the analyzer and its corresponding documentation. It served as an additional reference point, complementing the quality division's protocol numbering system.
A Provincial POCT Equipment Master Inventory List was prepared and regularly updated by the Eastern Zone POCT Coordinator. This centralized inventory included comprehensive details for all POCT equipment installations across the province, supporting ongoing standardization and quality assurance efforts.

Documentation
Accession form and Pre-Move Check List
The Accession form and Pre-Move Check List were completed if analysers were shipped from Eastern-Urban zone to other zones, and records were maintained by POCT coordinator.

Equipment Transfer policy
A robust policy was developed to align with the equipment shipment from one site to another including indemnity clause for vendors to bear the losses if there is any damage to the equipment during shipment of analysers.

Verification Protocol
Protocols for Primary device verification and secondary verification were prepared and signed off before starting verification processes.

Standard Operating Procedure and Quick Reference Document
These documents were prepared, reviewed and signed off before the placement of equipment at the final site.

Verification of analysers
Verification of these devices was completed in two phases.

Primary Verification
Operator training was conducted for all team members involved in the primary verification process. A comprehensive primary verification protocol was developed to guide the procedure. Two reference analysers were selected for in-depth validation at the primary verification site at SCH. The lab-based analysers, already in use at the site, were designated as the reference instrument for comparison. The primary verification process included a series of critical performance evaluations such as:

- Precision (Repeatability and Reproducibility)
- Linearity or accuracy studies
- Clinical correlation or method comparison
- Reference range verification, including pediatric populations
- Interference studies, wherever applicable

Once the analysers demonstrated acceptable performance, these were designated as the reference analysers for all subsequent secondary verification procedures conducted at the final installation sites e.g. Piccolo Chemistry analysers at 23 remotes sites were compared with primary validated Piccolo analyzer installed at St Clare's Mercy Hospital, the primary site of verification. Alternatively, if specific site was previously equipped with any other lab-based analyzer, then the oncoming device was compared with that site specific device e.g. Coaguchek devices were compared to site specific coagulation analysers for PT-INR testing. Process adjustments were made as necessary depending on the instrument model and sample type requirements, ensuring flexibility while maintaining scientific rigor.

Secondary verification Phase
Instruments were shipped to final sites. Operator's training was provided. Secondary verification studies were designed according to the site and analyzer requirement. Reproducibility was checked for all analysers and accuracy of these secondary devices was tested either using linearity material or clinical samples. Analysers were put into use once verification was accepted.

Connectivity
The implementation of POCT technologies included connectivity solutions that standardize data management (e.g.

through common middleware) across all provincial sites of NL to reduce the costs of infrastructure, upgrades, and ongoing operational expenses.

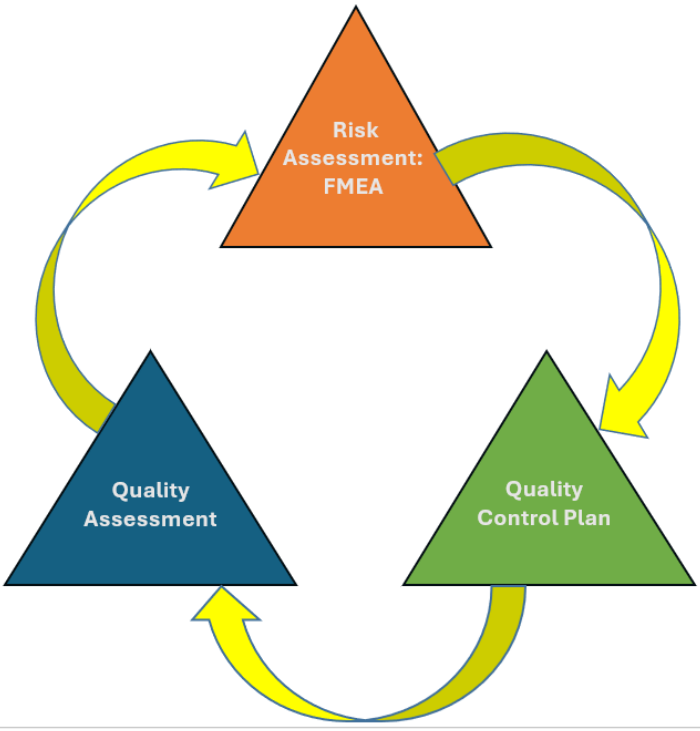
Electronic Tracking of Inventory of reagents and equipment
Due to nonavailability of centralized electronic tracking system for POCT reagent and equipment, tracking of calibrators, validated reagent lot numbers, cartridges and control material was done through shared Microsoft excel folder and information regarding availability of validated stock was transferred to the provincial users through a common group email. Master Inventory List was saved on this shared folder, and access was provided to all authorized users. Shared database and common email are helpful to track lot numbers of validated POCT reagents, calibrators, cartridges and controls and reduce the cost of validating same lot number of reagents multiple times and at multiple locations.

Instrument connectivity for patient result tracking
POCT results were electronically transmitted to the medical record to improve data quality and to support safe patient care. It is important that POCT results were reported in a manner which clearly distinguishes them from tests done using traditional central laboratory equipment e.g. prefix POC was added with name of each test for easy differentiation. Instruments are currently connected to the hospital information system network of EPIC. The EPIC system is connected to Navify, the POCT middleware.

Developing the IQCP for NLHS-POCT program
The Individualized Quality Control Plan (IQCP) is an inherent part of the quality improvement initiative. The traditional quality control (QC) processes have largely been replaced by IQCP [16]. CLSI in its recent guideline EP-23 A, advocated inclusion of individualized quality control plan [17]. According to this guideline, IQCP processes should be divided into 3 phases: pre-analytical, analytical and post-analytical. Each phase should contain appropriate control processes at the points where there is a possibility of error, to either prevent or monitor them and to take corrective action.
Keeping these requirements in view, we developed individualized plan for our POCT program where 3 key steps of IQCP were included (Figure 3).

1. **Risk assessment:** Risk assessment using failure mode and effects analysis (FMEA) tool was used to identify weaknesses of quality assurance process of every phase and plan mitigation.
2. **Quality Control Plan:** The plan was prepared to identify the processes where there was a maximum possibility of committing errors e.g. in analytical phase.
3. **Quality Assessment:** The monitoring of the IQCP included performance of testing personnel, testing environment, specimens, reagents, and test system [18].

Figure 3: Individualized Quality Control Plan for NLHS-POCT program.



IQCP for NLHS POCT Program

IQCP included following key points:

1. It was important to understand all aspects related to development of procedure, policy, procurement, accreditation, evaluation and validation, personnel education and training, quality assurance, and utilization management of POCT technologies.
2. The program collaborates with other provincial Pathology and Laboratory Medicine divisions for implementing and maintaining POCT service delivery models. We used cutting edge POCT technologies across all hospitals, long term care, and rural health facilities.
3. The multidisciplinary committee, named "POCT advisory council" was established mainly to authorize and approve implementation of POC tests across the province. Members of committee are from various disciplines of NLHS and carry expertise in their field. Selection of tests are done through voting on application submitted by clinical program. Members vote in favor or against, largely depends on appropriate requirements of clinical program, quality of the tests, competency and commitment of users to follow policies and maintain quality management system set up by the program.
4. The Division developed standardized policies, procedures, quality assurance protocols, and processes for the province and ensured that appropriate information is collected to assess the clinical justification, desired patient outcomes, financial impacts, infrastructure support and educational/ competency requirements of the specific point of care test.
5. Risk assessment for each project implementation was done by members of quality division. Risk assessment was carried out using FMEA technique to calculate the high-risk priority number of each process and analyze improvement strategy for weak processes.
6. It is mandatory for each POCT operator to comply with established policies and procedures regarding appropriate sample collection, quality control and reagent checks, basic maintenance (if applicable) and cleaning protocols for equipment, test performance, reporting, follow-up of all critical results, documentation and performance of proficiency testing. These professionals must perform POCT only after getting trained, and completion of subsequent recertification at prescribed intervals.
7. Only certified users who demonstrated competence in specimen collection and performing a specific POC test were permitted to operate the devices and report the results.
8. Appropriate internal QC, electronic QC or external QC procedures were designed.
9. Related proficiency test scheme was selected for each test. If appropriate commercial proficiency test scheme was not available, then alternate methods like inter-lab comparison or split sampling procedures were introduced.

Setting up internal quality control frequency

A designated QC procedure was developed for each POCT analyzer. The QC frequency for each system was developed based on recommendations given by Norwegian Quality Improvement of Laboratory Examination (NOKLUS) QC scoring system for POC tests as depicted in figure 4 [19]. POC analysers were of different types e.g. multi-test cartridge, single strip test, bench top analyzer etc.

The Scoring System

QC was used each time if

a) A new batch of reagent or test was used,

b) After unexpected test results, and

c) After maintenance and instrument repair.

In addition to this mandatory QC performance, the frequency of QC was established based on a scoring system.

The general principle of the scoring system was to mitigate any high risk of harm to the patients; therefore, QC was performed at an appropriate frequency.

Frequency Calculation

For each POCT device, factors A, B, and C were given points as depicted in scheme (Figure 4) to calculate the total score by adding each individual factor score. The total score determines QC frequency for each POC test. For example, if for ALT, total score (A+B+C) is between 7 - 9, then QC frequency should be weekly.

In addition to this score, frequency is tweaked depending on the number of patient samples analyzed over a specific period (Factor D). We adopted this scoring system to decide performance of QC at various intervals for each point of care device (19).

Scoring was given as below:

Analyte Score: was given based on the risk of harm caused to the patient if any error occurred. There were two categories:

Moderate risk: 2

High risk: 4

Device Score: Device score was given depending on multiple types of POC device which could be scored as follows:

Strip test: 1

Strip test with automated readers like Glucometer: 2

Single Cartridge based analyzer: 3

Small Bench top models like blood cell counter: 4

Ease of Use: User Friendliness of a device was scored as follows

Easy: 1

Moderate: 2

Difficult A1C reader: 3

Blood cell counter: 4

Changes in QC frequency depending on the work volume: number of samples tested over a specific time.

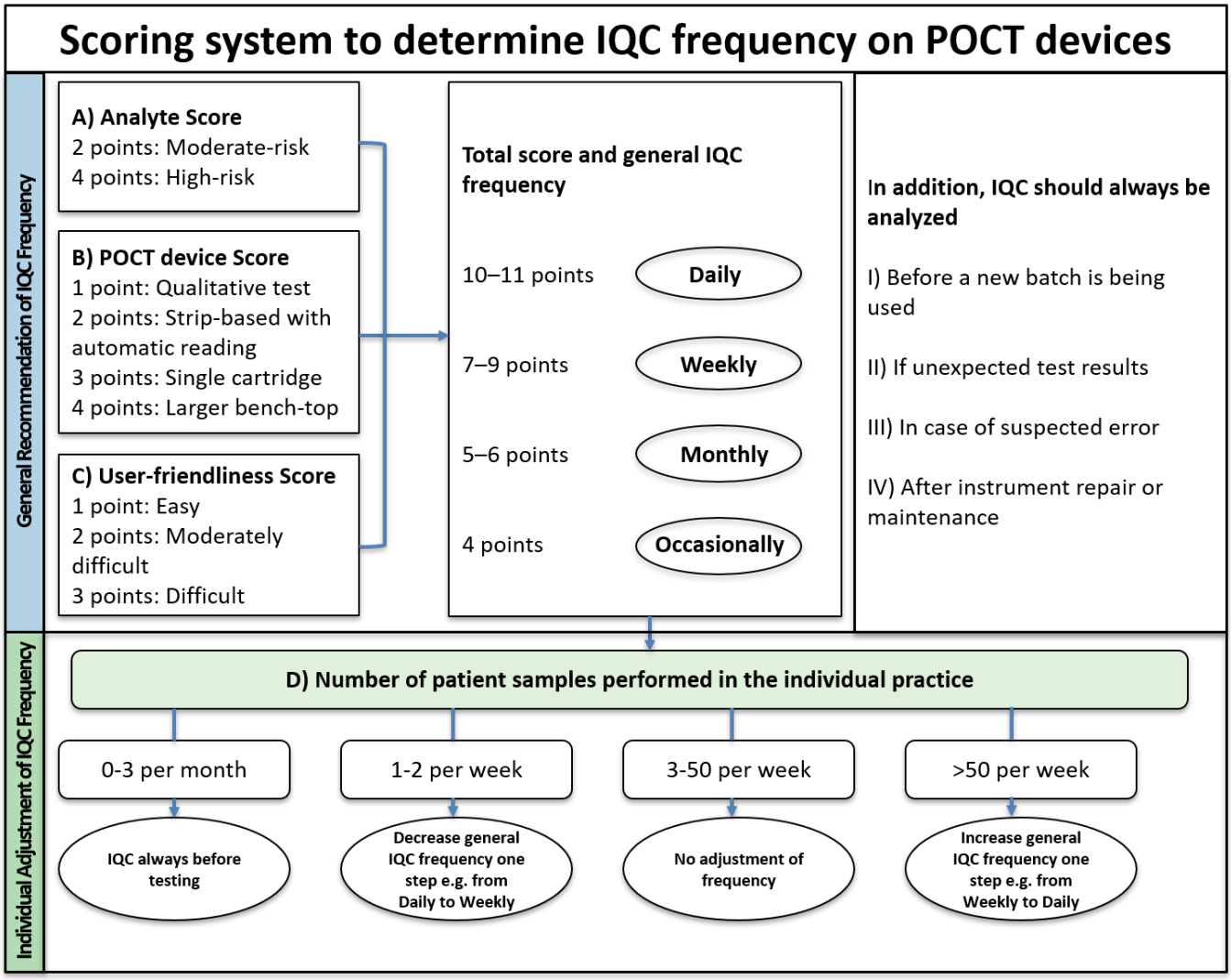
If total score was between

- 11 points; QC was performed daily, every day that the test was in use

- 9 points; QC should be performed weekly
- 6 points; QC should be performed Monthly
- 4 points; QC should be performed occasionally

Depending on the work volume, this frequency was modified (Figure 4).

Figure 4: Quality control frequency for all Point of Care devices was developed based on NOKLUS recommendation of Daily Quality Control Frequency scoring scheme. (Reused with permission from authors and journal editors, Gidske G, Sandberg S, Fossum AL, et al. Point-of-care testing in primary healthcare: a scoring system to determine the frequency of performing internal quality control. Clin Chem Lab Med. 2022; 60(5):740-747 [19]).



Program effectiveness evaluation

Utilization management and cost effectiveness

Utilization management interventions to evaluate the medical necessity, efficiency, and appropriateness of usage of POCT is an important step towards increasing work efficiency and cost effectivity. Scope of POCT is expanding beyond the clinics today. Tests are being carried out deep in sea and at the heights of Mount Everest [20]. Usage of POCT especially by non-laboratory users outside the central laboratory, in acute care settings, physician's offices, pharmacies, ambulances, and outpatient clinics presents a unique challenge for any POCT program to assess the utilization [21].

Currently, there are no established structured processes that will routinely monitor and manage utilization of POCT within our health system i.e. Newfoundland and Labrador health services. The preparation of structured utilization monitoring tools to assess financial implications and cost effectiveness is under development. To develop a map of cost effectiveness for our program, intervention techniques like prior authorization, concurrent review, control costs, ensured quality care, audit and manage resource utilization need to be included [22]. These techniques will act as a bridge between users and the providers. We expect that the developed metrics will be helpful in analyzing the cost effectiveness of the project against rising cost, inflation, supply chain bottlenecks, overutilization and underutilization of tests due to poor understanding of caregivers.

Developing structured metrics for any program beyond laboratory is a daunting task. The biggest challenge is to enforce accountability for users who are repeated offenders for overutilizing or underutilizing POCT resources to obtain quick patient results. The users are in large number and have different qualifications and skillsets. Each user works for a different clinical program and reaching out to them individually is impossible hence we have developed a standardized process of sending out memos and opted to include updated information on the intranet and pulse web. However, surprisingly not many users spend time reading the required essential information on these platforms.

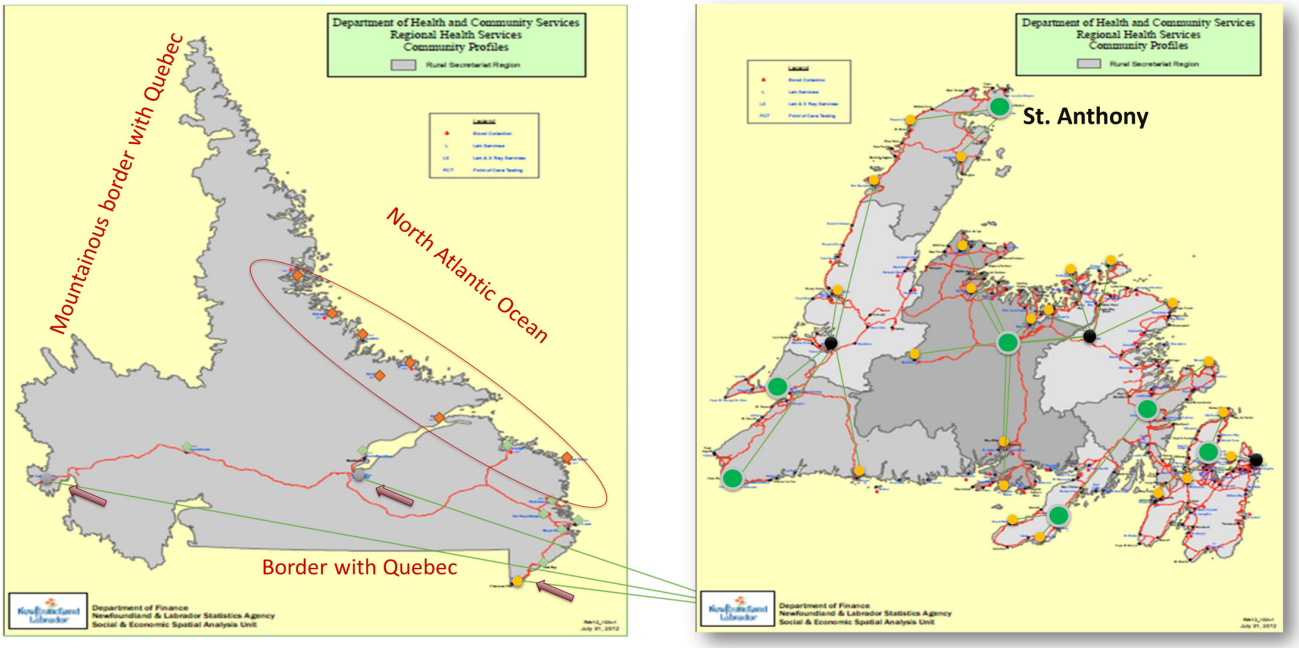
Audit is an integral part of a healthcare system to monitor utilization of resources. Auditing any hospital unit, to reduce the overutilization involves systematic approach of reviewing test orders against clinical necessity, focusing on reducing unnecessary, repetitive, or low-value testing, which is estimated to account for 20% to 40% of laboratory tests [23]. Use of evidence-based clinical guidelines is a helpful tool in determining if multiple and repeated test requests are necessary for clinical interpretation and diagnosis of diseases or not. The audits are planned to be carried out by some authorized and certified auditors working within NLHS. Quality Officers and Operation Managers having provincial POCT responsibilities will be designated with the jobs of reviewing and auditing POCT utilization on a regular basis. Utilization audits will be conducted at defined intervals to assess the cost efficiency of the program.

Effectiveness and sustainability of any healthcare project is primarily affected by the increasing financial burden. Overutilization or underutilization greatly affects expenditure. The cost of consumables (test kits, QC materials, etc.) is significantly higher for POC tests compared to traditional laboratory tests [24]. Simeon et al. [24] also reported that costs per POC test vary substantially, depending on the number of tests performed on an instrument.

Removing inefficiencies and equipment sharing between clinics appears to help decrease costs in smaller and more remote clinics. These clinics are unable to absorb the costs of equipment individually.

There is an additional cost of transportation of samples from remote clinics to centralized labs. NLHS POCT program covers a vast majority of remote locations separated from one another by a greater distance (Figure 5) with no immediate laboratory support, therefore, assessment of factors responsible for sustainability of these remote POCT sites becomes mandatory. Simeon et al did not evaluate costs related to increasing distance from centralized laboratories which include an important cost factor of transportation of samples from remote clinics to lab [24].

Figure 5: Location of POCT and laboratory sites operated by Newfoundland and Labrador Health Services. Yellow Circles: Location of small community health centers transitioned to become POCT sites. Orange Circles: Isolated coastal community sites changed to POCT. Green and Grey Circles: Larger community hospitals with limited-service in-house laboratories. Black Circles: Indicate sites with larger full service (Pathology, Microbiology, Clinical Chemistry, and Hematology) laboratories.



Another important aspect of cost saving is to choose the tests wisely and appropriate to the clinical need to reduce burden on any healthcare program. Our recent article on "Choosing Wisely Canada for POCT" describes recommendations on choosing wisely using POCT in Acute care settings. This guidance document is a useful tool for POCT users working in emergency and acute care hospitals for improved patient care through strong diagnostic integrity, augmented safeguards to patients, and optimization of related healthcare resources [25].

Key Performance Indicators or Quality Indicators

Our POCT program is still emerging. Quality improvement processes are continually assessed for their effectiveness through the key performance indicators or quality indicators. These quality indicators help in improving the processes. These are designed to identify the criteria to monitor, measure and analyze the processes [26]. NLHS POCT program has established a strong provincial quality indicator special interest group named as "Newfoundland & Labrador Quality Indicator Interest Group (NLPQIIG)". Function of this group is to develop standardized provincial POCT QIs. The group, which is chaired by Medical Scientific Lead for POCT, includes all key staff members of POC and quality division. The group holds monthly brainstorming meetings and is working towards developing quality indicators to standardize the processes across 23 POCT sites.

Two most effective quality indicators have been selected and implemented to monitor CQI and CPI. Choice of these two indicators aligns with recommendations given by our Canadian POCT QI special interest group of CSCC, which is coauthored by lead author of this manuscript [27]. These indicators are,

1. Percentage of invalid patient identification or generic ID use for performing tests on glucose meters

and

2. Critical glucose results follow up.

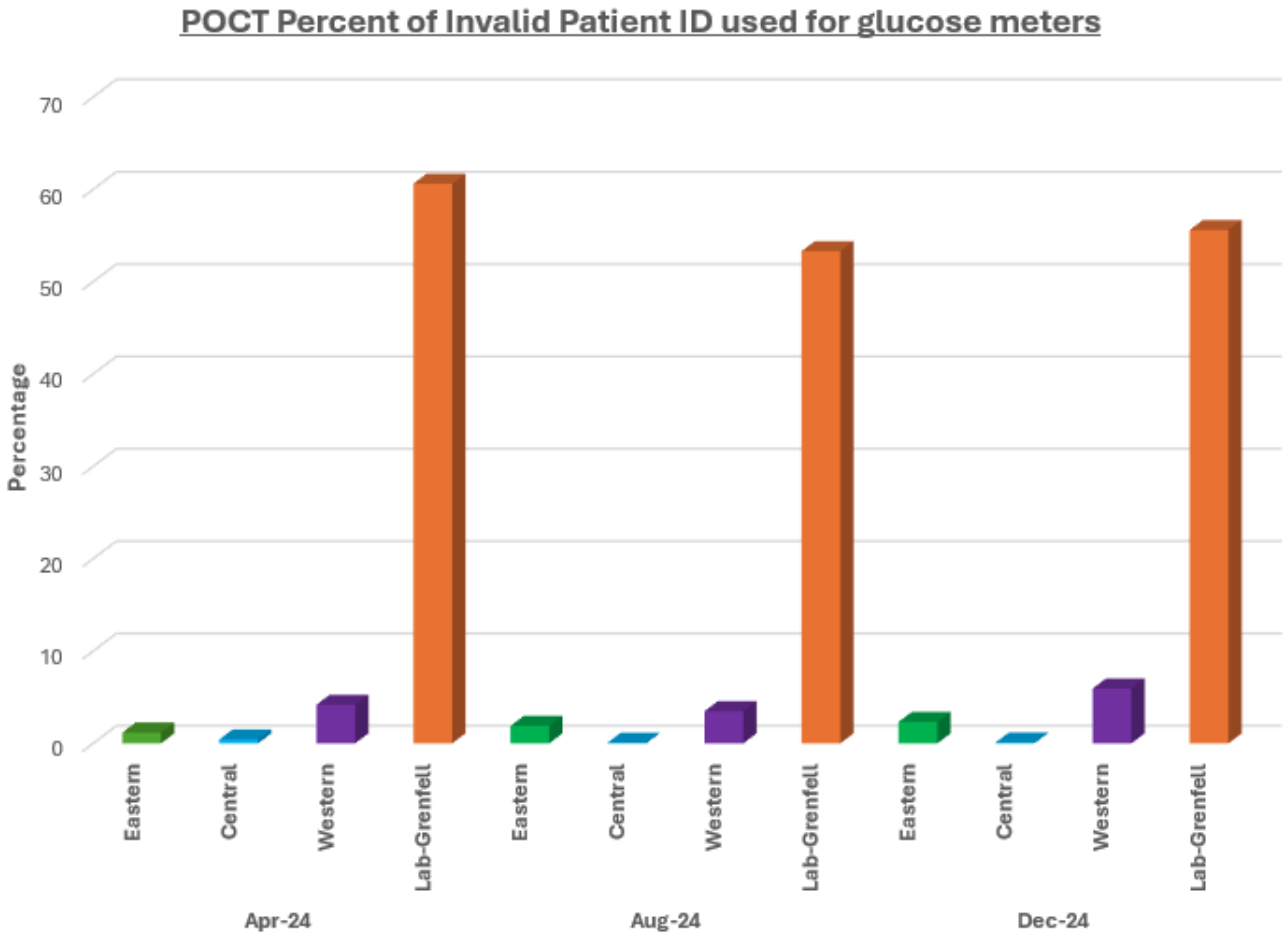
The standardization of process of quality indicator data collection across our program, including 23 sites is in initial stages. Implementation of one process across these many sites is difficult due to several factors like limited available resources, staffing challenges, different information system across all zones (Meditech) which has recently been replaced by new EPIC system, multiple certified glucometer users, and different laboratory levels of their competency and commitment towards following recommended guidelines. Continuous quality and process improvement (CQI and CPI) procedures will be used across the province after full implementation of EPIC hospital information system replacing meditech.

Percentage of invalid patient identification or generic ID use for performing tests on glucose meters

Figure 6 shows comparative data (expressed in percent) for the use of Invalid Patient ID (e.g.999999) for testing glucose on glucose meters across all 4 zones. Data collected is from the three quarters of year 2024. Eastern zone data includes both rural and urban eastern zone sites. The highest use of generic ids was seen in Labrador-Grenfell health zone. The health

zone is catering to multiple remote indigenous communities and has the highest number of isolated remote POCT clinics. These clinics are located at greater distances from major acute care centres and operated by locum nurses and physicians. These staff members are not available for follow up calls to get the valid patient identification number against used generic identification numbers. The users did not complete the required documentation to match the generic ids with actual patient identification numbers.

Figure 6: Percent use of invalid patient ID for glucose meters across four different healthcare zones (Eastern, Central, Western and Labrador-Grenfell), NLHS in 3rd quarter of year 2024. Labrador-Grenfell (Lab-Grenfell) healthcare zone had highest number of invalid or generic id usage for testing patient samples on glucose meters. POC program did not receive completed documentation for actual patient identification number for the generic numbers used (e.g.999999) even after repeated requests. Lowest use was in central healthcare zone.



Critical Glucose Result follow up

"Critical glucose results follow up" is an important quality indicator for POCT. This indicator is focused on reducing patient safety risks. The extremely high (> 25 mmols /L) or very low (<2.5 mmols /L) blood glucose results must be rechecked for confirmation before releasing the results on patient's medical records. Our POCT program has a policy for "Retesting Critical Glucose" results within 10 minutes of first result, either on another POCT meter or by sending a venous blood draw to the central laboratory. The confirmatory patient results must be documented in patient charts. Nonlaboratory users, mainly nurses, are required to complete the task. The process has been developed to reduce the patient risk.

We do not have recent data on this QI from all 4 zones. Only two sites provided us the data on critical glucose retesting which showed 75-80% compliance on policy implementation (data not presented).

It is planned to collect more robust and reliable annual data by the end of April 2027 (post EPIC) from all 4 zones. This data will clarify and effectively assess implementation of Lean processes across NLHS POCT program.

Turn Around Time (TAT)

TAT calculation is not a useful indicator as POCT uses a "test and treat" model and these devices are used only to perform stat and urgent tests. The results obtained from these devices

are auto verified in the EPIC and immediately available for clinical use.
Collection of TAT data for the samples sent from these clinics to hub lab for analysis is not possible as several sites are located as standalone and remote clinics. Transportation of samples from these clinics to hub labs is affected by multiple variables like adverse weather conditions, difficult terrains and accessibility for sample pickups, different transportation modes, bad road conditions, nonavailability of 24/7 transportation as some sites are only connected by sea route.
Quality indicator data collection previously used did not provide any reliable data as these procedures were not uniform and were different for each regional healthcare authority. NLHS was divided into 4 separate healthcare authorities until 2025. These authorities worked as independent entities and had customized processes suitable for them. Challenges were encountered in implementation of uniform procedures of QI data collection due to lack of a single robust hospital information system (Meditech) throughout these 4 authorities. The old Meditech builds were customized as per each authority's requirements. Staff members like POCT coordinators and other quality team members working for these authorities were more focused on collecting QI data only to fulfil the ACD regulatory requirements and to present the data to auditors during annual audits. They did not work towards provincial standardization of QI collection.
There is a great scope post EPIC for implementation of continuous quality and process improvement procedures in future. The standardization will be much easier and robust after full implementation of EPIC Hospital Information System across the province.

Discussion

POCT has emerged a transformative approach in rural and remote healthcare delivery, offering significant advantages over conventional laboratory testing, particularly because it eliminates the need for dedicated laboratory infrastructure. It enhances mobility for both staff and patients, provides portability across remote and community-based locations, and ensures faster turnaround times for diagnostic results. Clinicians desire a rapid, reliable and efficient service delivered at low cost; of these characteristics, timeliness is perhaps the most important. Timeliness remains a critical determinant of clinical utility, as clinicians often prioritize rapid availability of results to guide immediate decision-making. The integration of Lean Six Sigma methodologies into healthcare system has further enhanced operational efficiency across several clinical environments, including emergency departments, cardiology, surgery, intensive care units, pediatrics, pharmacy, radiology, and laboratories [6-10]. The application of these methodologies within POCT program development represents a strategic advancement toward standardized, efficient, and scalable diagnostic services. Over years several regulatory changes have been implemented in the building of POCT devices by diagnostic companies, to ensure compliance with quality assurance processes like the centralized laboratory service. In Canada, all near-patient, In Vitro Diagnostic Devices (IVDDs), including most POCT devices, are classified as Class III and must undergo pre-market evaluation by Health Canada to obtain a Medical Device License [28]. All POCT equipment/ devices require regular quality control testing to be performed at defined intervals to ensure that the results are accurate and reliable and the device's technical performance is within the expected performance checks [29].
More so, POCT represents a "Test and Treat" model, which has successfully been used in Cervical Cancer Prevention Program of Papua New Guinea [5]. This use of Xpert HPV assay on self-collected vaginal specimens with same-day curative cervical ablation had proved best 'Test and Treat' tool for these patients. Any visual inspection of the cervix with acetic acid (VIA) had shown poor performance when used as part of screen and treat strategies for cervical cancer prevention [5].
In our previous review [30], we highlighted the unique leadership and logistical challenges associated with implementing POCT programs across large geographic areas. These difficulties were evident not only in low- and middle-income regions of Africa but also in remote parts of developed countries like Canada. In rural locations of NL, challenges included the lack of trained personnel, limited infrastructure, and poor transportation access, all of which contributed to inefficiencies, increased waste, and delayed test results. To address these challenges, we employed RPIW, a Lean Six Sigma tool, designed to eliminate operational waste, streamline workflows, enhanced team collaboration, standardized procedures, and reduced turnaround times [12-13, 31]. RPIW has been extensively used in several sectors of healthcare system. For instance, RPIW was used to develop and implement a playbook employed for automated patient-facing assessment with efficient collection and communication of patient information in Veteran Health Administration [12]. RPIW has also been used in optimizing hierarchical condition category-risk adjustment factor management in population health [32].
The primary objective of adopting the RPIW framework was to lay the foundation for a strong, standardized POCT program across rural and remote communities in Newfoundland and Labrador. As a foundational step, a multidisciplinary team was assembled to assess and guide the implementation of key Lean Six Sigma components within the provincial POCT initiative. The RPIW model proved adaptable to NL's needs, with structured planning phases allowing diverse stakeholders to align on shared objectives. The initial intensive workshop enabled real-time problem-solving, process mapping, and future-state modeling, ensuring that waste (Muda) was minimized and workflows streamlined. This structured approach provided a repeatable model for future initiatives within the POCT division and potentially other laboratory units. The creation of standardized documentation, particularly

the PEUIN system, brought much-needed transparency and traceability to analyzer installations. This supported improved regulatory compliance and audit readiness, while facilitating centralized inventory control. The use of unique identifiers, in conjunction with verification protocols, allows clear documentation alignment and performance tracking, mitigating risks related to manual data entry and fragmented tracking systems.

The two-phase verification process adopted in this program represents a robust approach ensuring high analytical integrity. Conducting primary verification centrally, allowed for thorough analytical assessments using reference instruments, while secondary verification at local sites ensured reproducibility in real-world conditions. This dual-step model enhanced the reliability and comparability of POCT results across sites, contributing to better clinical decision-making.

The incorporation of the IQCP as a core part of the quality framework signified a shift from traditional QC practices to a more risk-based, device-specific approach. This approach is consistent with international models such as the Norwegian Quality Improvement of Laboratory Examination (NOKLUS). The integration of a scoring system based on NOKLUS recommendations allowed POCT teams to adjust QC frequency dynamically, ensuring appropriate balance between patient safety and operational efficiency [19]. This system emphasized patient-centered care by recognizing higher-risk analytes and devices and ensuring more frequent QC as needed. For example, analysers used for high-risk tests such as blood gases or glucose in critical care settings received daily QC, whereas low-risk tests on stable populations could be monitored less frequently without compromising safety. For almost two and half decades, NOKLUS had focused on POCT in primary healthcare laboratories. The NOKLUS quality system used different tools to obtain harmonization and improvement like external quality assessment for the pre-examination, examination and post examination phase to monitor the harmonization process and to identify areas that need improvement and harmonization. This program also included manufacturer-independent evaluations of the analytical quality and user-friendliness of POC instruments. Close interactions and follow-up of the participants through site visits, courses, training and guidance is a part of NOKLUS program [33]. NOKLUS was established in 1992 as the Norwegian Quality Improvement of Primary Care Laboratories, with the purpose of improving analytical quality of the POC laboratory examinations. This is now re-named as Norwegian Quality Improvement of Laboratory Examination and has more than 3100 voluntary participants [34-35].

An earlier review article [31] from our group on setting up the provincial POCT network for rural communities of Newfoundland and Labrador described expected benefits and challenges for setting up this program in remote locations. Another review article from an Australian group has described advantages for setting up POCT program for largely replacing conventional laboratory setup due to large investments and capital funds required [36].These authors described POCT as a beneficial approach over conventional design of laboratory to avoid delays in clinical decision making due to late availability of the results from lab, increased costs for purchase and maintenance of equipment, staff training, connectivity to the laboratory information system (LIS), quality control (QC) and external quality assurance (EQA) procedures, some essential steps needed to fulfil the regulatory compliance requirements of ISO 15189: 2022 [36].

An important aspect of POCT, which has always received very little attention, is utilization management. POCT usage is vulnerable to the same risks for overuse and underuse as other laboratory tests. The cost of consumables (test kits, QC materials, etc.) tend to be significantly higher for POCT compared to traditional laboratory tests. It is a common belief that the cost of point of care tests is significantly higher than the traditional laboratory testing, however, we did not come across any peer reviewed journal article describing the overall cost effectiveness of the POCT over traditional laboratory testing. However, there are a few reports available in the literature evaluating cost effectiveness of different POC tests compared to traditional laboratory settings. The data from a study which was completed in the year 1999 has shown that by replacing central laboratory testing of blood samples both at an accident and emergency (A&E) department and at an intensive therapy unit (ITU) by POCT decreased total hospital cost from £132,630 to £93,050 [37]. El-Osta et al have developed a mathematical model of cost-minimization analysis to investigate the potential cost-saving use of POCT in delivering NHS Health Check in the primary care setting [38]. The National Health Service (NHS) Health Check program is England's national CVD primary prevention program launched by the Department of Health in April 2009 and people aged 40–74 years and free of vascular disease diagnosis were registered for an NHS Health Check. The program used information from the Health Check to calculate a CVD risk score. The score is based on results of blood sugar (glucose or HbA1c) and total cholesterol levels. These samples are tested either by laboratory or POCT equipment. GPs relying on laboratories first schedule phlebotomy tests and afterwards conduct the full NHS Health Check using glucose and cholesterol readings to calculate and communicate CVD risk. Lab based pathways require patients to book several appointments to get their CVD score and complete the NHS Health Check. The NHS Health Check program delivered using POCT equipment obtains immediate total cholesterol and glucose/HbA1c results using blood samples taken from 'finger pricking' patients. The use of POCT facilitates communication of CVD risk scores to patients on a single visit. According to the results of their mathematical model, the cost savings to the NHS associated with POCT were estimated at £29 per 100 patients.

In 2020, Wong et al [39] published a review on health economics of point of care testing worldwide. According to

these investigators, POCT could potentially reduce costs, enhance workflow efficiency and improve patient care by enabling more prompt diagnosis and treatment decisions, cost-saving technology solutions and wireless connectivity. They did find some studies who have reported higher rates for implementation and sustainability of POCT [40]. According to these investigators, POCT design, outcomes and economics are strongly linked to economic status. Ultrasound-based POCT is the most implemented type of POCT in developing countries, while biomarker/lab test-based POCT is prevalent in developed countries. Biegler et al have published a report on use of POC system in emergency departments for rapid analysis of blood samples resulting in a substantial saving of $111 per patient, with total savings from Canada of $7,350,000 per year in the health care setting [41].

POCT instruments are used at various locations in hospitals, health care institutions, ambulances, at physician offices, in flights and ships. Providing connectivity for hundreds of POC devices with thousands of users become a difficult task for any provincial program. Network middleware serves as a critical digital bridge between these POC analysers like glucose meters, blood gas analysers, chemistry or CBC analysers and the centralized information systems like the Laboratory Information System (LIS) or Electronic Medical Record (EMR).The use of information systems directly interfaced to the point of care devices through a middleware ensure that results obtained from these analysers are being transmitted without manual interventions. This helps in reducing the transcription error [42].

There are multiple network providers available in the market to make POCT middleware available to the users. These networks are all HL7 compatible which is used for communication between the middleware and the Laboratory Information System, enabling the integration of data from multiple devices into a single Electronic Health Record system [42]

Limitations and future directions

Although the RPIW framework is successfully implemented, certain limitations remain. The full transition to the Employees Participating in Change (EPIC) information system introduced temporary workflow disruptions. Training provided by the EPIC team to POCT staff members is adequate just to introduce the staff to the new system, however, familiarization of users to the new HIS will take months. A robust structured framework to evaluate utilization of POCT and implement uniform processes of pulling key performance indicator data is currently under development. The proposed provincial materials management system for reagent lot tracking represents a significant advancement, its development and adoption is currently in progress. Continuous monitoring, stakeholder engagement, and responsiveness to evolving clinical and technological needs will be essential to sustain the gains achieved through this initiative.

Abbreviations

ACD: Accreditation Canada Diagnostics
ALT: Alanine Transaminase
CBC: Complete Blood Counts
CLSI: Clinical & Laboratory Standards Institute
CPI: Continuous Process Improvement
CQI: Continuous Quality Improvement
CSCC: Canadian Society of Clinical Chemistry
CVD: Cardiovascular Disease
DMADV: Define, Measure, Analyze, Design and Verify
EPIC: Employees Participating in Change
FMEA: Failure Mode and Effects Analysis
GLP: Good Laboratory Practice
GPs: General Physicians
HIS: Hospital Information System
HL7: Health Level 7
H Sc: Health Sciences Centre
HPV: Human Papilloma Virus
Ids: Identification numbers
IQCP: Individualized Quality Control Plan
KPI: Key Performance Indicator
NHS: National Health Service
NLHS: Newfoundland and Labrador Health Services
NLHS-POCT: Newfoundland and Labrador Health Services-Point of Care Testing
NOKLUS: Norwegian Quality Improvement of Laboratory Examinations
PDSAI: Plan, Do, Study, Act and Improve
PEUIN: Provincial Equipment Unique Identification Number
POCT: Point of Care Testing
PT-INR: Prothrombin Time-International Normalized Ratio
QC: Quality Control
QI: Quality Indicator
RPIW: Rapid Process Improvement Workshop
SCH: St. Clare's Mercy Hospital
SOP: Standard Operating Procedure
TAT: Turn Around Time

Funding resources

This research did not receive any specific grant from funding agencies in the public, commercial, and not for profit sectors.

Conflict of interest and ethical approval

There is no conflict of interest for any author. Article is based on a management aspect of POCT program; therefore, study does not require ethical approval.

References

1. Price CP. Point of care testing. BMJ. 2001; 322(7297):1285-1288. DOI: 10.1136/bmj.322.7297.1285.
2. Chaisirin W, Wongkrajang P, Thoesam T, et al. Role of Point-of-Care Testing in Reducing Time to Treatment Decision-Making in Urgency Patients: A Randomized Controlled Trial. West J Emerg Med. 2020; 21(2):404-410.

DOI: 10.5811/westjem.2019.10.43655.
3. Shaw JLV. Practical challenges related to point of care testing. Pract Lab Med. 2015; 9(4):22-29. DOI: 10.1016/j.plabm.2015.12.002
4. Zhaowei Z, Peng M, Rajib A, et al. Advanced Point-of-Care Testing Technologies for Human Acute Respiratory Virus Detection. Adv Mater. 2022; 34(1):e2103646. DOI: 10.1002/adma.202103646.
5. Qi S, Xindi S, Ziyi D et al. Point-of-care testing detection methods for COVID-19. Lab Chip. 2021; 21(9):1634-1660. DOI: 10.1039/d0lc01156h.
6. Zimmermann GDS, Siqueira LD, Bohomol E. Lean Six Sigma methodology application in healthcare settings: an integrative review. Rev Bras Enferm. 2020; 73(Suppl 5): e20190861. DOI: 10.1590/0034-7167-2019-0861.
7. de Barros LB, Bassi L de C, Caldas LP, et al. Lean healthcare tools for processes evaluation: an integrative review. Int J Environ Res Public Health. 2021; 18: 7389. DOI: 10.3390/ijerph18147389.
8. Peimbert-Garcıa RE. Analysis and evaluation of reviews on Lean and Six Sigma in health care. Qual Manag Health Care. 2019; 28:229–236. DOI: 10.1097/QMH.0000000000000226.
9. Hern´andez-Lara AB, S´anchez-Rebull MV, Nin~erola A. Six sigma in health literature, what matters? Int J Environ Res Public Health. 2021; 18:8795. DOI: 10.3390/ijerph18168795.
10. Rotter T, Plishka C, Lawal A, et al. What is lean management in health care? Development of an operational definition for a Cochrane systematic review. Eval Health Prof. 2019; 42:366–390. DOI: 10.1177/0163278718756992.
11. Nin~erola A, S´anchez-Rebull MV, Hern´andez-Lara AB. Quality improvement in healthcare: six sigma systematic review. Health Policy. 2020; 124:438–445. DOI: 10.1016/j.healthpol.2020.01.002.
12. Pittman JOE, Rabin B, Almklov E, et al. Adaptation of a quality improvement approach to implement eScreening in VHA healthcare settings: innovative use of the Lean Six Sigma Rapid Process Improvement Workshop. Implement Sci Commun. 2021; 2:37. DOI: 10.1186/s43058-021-00132-x.
13. De Koning H, Verver JP, van den Heuvel J, Bisgaard S, Does RJ. Lean six sigma in healthcare. J Healthc Qual. 2006; 28(2):4–11. DOI: 10.1111/j.1945-1474.2006.tb00596.x.
14. Sari N, Rotter T, Goodridge D, Harrison L, Kinsman L. An economic analysis of a system wide Lean approach: cost estimations for the implementation of Lean in the Saskatchewan healthcare system for 2012-2014. BMC Health Serv Res. 2017;17(1):523. DOI: 10.1186/s12913-017-2477-8.
15. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. BMJ Qual Saf. 2014;23(4):290–8. DOI: 10.1136/bmjqs-2013-001862.
16. Khan AI. Best Laboratory Practices Regarding POCT in Different Settings (Hospital and Outside the Hospital). EJIFCC. 2021; 32(2):124-130.
17. Nichols JH. Laboratory quality control based on risk management. Ann Saudi Med. 2011; (3):223-228. DOI: 10.4103/0256-4947.81526.
18. https://www.cms.gov/medicare/quality/clinical-laboratory-improvement-amendments/quality-control
19. Gidske G, Sandberg S, Fossum AL, et al. Point-of-care testing in primary healthcare: a scoring system to determine the frequency of performing internal quality control. Clin Chem Lab Med. 2022; 60(5):740-747. DOI: 10.1515/cclm-2021-1258.
20. Nichols JH. Utilizing Point-of-Care Testing to Optimize Patient Care. EJIFCC. 2021;32(2):140-144. PMID: 34421482; PMCID: PMC8343046.
21. Chokkalla AK, Recio BD, Devaraj S. Best Practices for Effective Management of Point of Care Testing. EJIFCC. 2023;34(3):245-249. PMID: 37868087; PMCID: PMC10588082
22. Board on Children, Youth, and Families; Institute of Medicine; National Research Council. Considerations in Applying Benefit-Cost Analysis to Preventive Interventions for Children, Youth, and Families: Workshop Summary. Washington (DC): National Academies Press (US); 2014;3. Assessing the Costs and Benefits of Interventions. Available from: https://www.ncbi.nlm.nih.gov/books/NBK219358/
23. Luigi D, Emilie C, Patrick M, et al. Interventions to improve appropriateness of laboratory testing in the intensive care unit: a narrative review, Annals of Intensive Care, 2024;14, 1(9),2110-5820, https://doi.org/10.1186/s13613-024-01244
24. Simeon K, Sharma M, Dorward J, et al. (2019) Comparative cost analysis of point-of-care versus laboratory based testing to initiate and monitor HIV treatment in South Africa. PLoS ONE 14(10): e0223669. https://doi.org/10.1371/journal.pone.0223669
25. Kaur S, Arnoldo S, Taher J, et al. Choosing Wisely Canada recommendations for point of care testing utilization in the acute care setting: Expert-based Delphi consensus from the Canadian Society of Clinical Chemists. Clin Biochem. 2026;142:111100. doi: 10.1016/j.clinbiochem.2026.111100. Epub 2026 Feb 13. PMID: 41692183.
26. Accreditation Agreement, Qmentum Canada, standards Point-of-Care Testing, January 2019 Ver. 14
27. Shaw JLV, Arnoldo S, Beach L, et al. Establishing quality indicators for point of care glucose testing: recommendations from the Canadian Society for Clinical Chemists Point of Care Testing and Quality

Indicators Special Interest Groups. Clin Chem Lab Med. 2023;61(7):1280-1287. doi: 10.1515/cclm-2023-0147. PMID: 37043622.
28. Community based point-of-care-testing (POCT) framework: A report of Public Health Canada, August 2025
29. College of Physician and surgeon of Alberta. Performance of Point-of-Care Testing in Unaccredited Settings, A Guideline for Non-laboratorians, December 2021.
30. Thakur V, Akerele OA, Brake N, et al. Use of a Lean Six Sigma approach to investigate excessive quality control (QC) material use and resulting costs. Clin Biochem. 2023; (112):53-60. DOI: 10.1016/j.clinbiochem.2022.12.001.
31. Randell EW, Thakur V. Leading POCT Networks: Operating POCT Programs Across Multiple Sites Involving Vast Geographical Areas and Rural Communities. EJIFCC. 2021; 32(2):179-189.
32. Benjamin LK, Meyer CB, Pan J, et al Optimizing Hierarchical Condition Category-Risk Adjustment Factor Management in Population Health Using Rapid Process Improvement Methods. Popul Health Manag. 2024;27(6):365-373. doi: 10.1089/pop.2024.0147.
33. Stavelin A, Sandberg S. Harmonization activities of Noklus - a quality improvement organization for point-of-care laboratory examinations. Clin Chem Lab Med. 201819;57(1):106-114. doi: 10.1515/cclm-2018-0061. PMID: 29768244
34. Sandberg S, Christensen NG. Quality control of laboratory analyses performed in general practice. Tidsskr Nor Laegeforen 1988;108:1888–90
35. Sandberg S, Christensen NG, Thue G, Lund PK. Performance of dry-chemistry instruments in primary health care. Scand J Clin Lab Invest. 1989;49:483–81.
36. Florkowski C, Don-Wauchope A, Gimenez N, et al. Point-of-care testing (POCT) and evidence-based laboratory medicine (EBLM) - does it leverage any advantage in clinical decision making? Crit Rev Clin Lab Sci. 2017;54(7-8):471-494. doi: 10.1080/10408363.2017.1399336. Epub 2017 Nov 23. PMID: 29169287.
37. Kendall JM, Bevan G and Clancy MJ. Point of care testing in the accident and emergency department: a cost analysis and exploration of financial incentives to use the technology within the hospital. J. Health Serv Res Policy, 1999;4,1-33
38. El-Osta A, Woringer M, Pizzo E, et al. Does point-of-care testing improve cost-effectiveness of the NHS Health Check program in the primary care setting? A cost-minimization analysis. BMJ Open. 2017;7(8): e015494. doi: 10.1136/bmjopen-2016-015494. PMID: 28814583; PMCID: PMC5724165
39. Wong HY, Marcu LG, Bezak E, Parange NA. Review of Health Economics of Point-of-Care Testing Worldwide and Its Efficacy of Implementation in the Primary Health Care Setting in Remote Australia. Risk Manag Health Policy. 2020;13:379-386. doi: 10.2147/RMHP.S247774. PMID: 32440241; PMCID: PMC7212773.
40. El Helali N, Habibi F, Azria E, et al. Point-of-care intrapartum group B streptococcus molecular screening: effectiveness and costs. Obstet Gynecol. 2019;133(2):276–281. doi: 10.1097/AOG.0000000000003057
41. Biegler N, McBeth PB, Tiruta C, et al. The feasibility of nurse practitioner-performed, telementored lung telesonography with remote physician guidance - 'a remote virtual mentor'. Crit Ultrasound J. 2013;5(1):5).
42. Erasmus R, Sahni S, El-Sharkawy R. Connectivity Strategies in Managing a POCT Service. EJIFCC. 2021;32(2):190-194. PMID: 34421487; PMCID: PMC8343042

Brief Report

Evaluation of NicCheck I Test Strips for the rapid detection of urine nicotine metabolites

Seema Khattri Bhandari[1], James H Nichols[1*]

[1]*Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA*

Article Info

[*]***Corresponding Author:***
James H Nichols
Vanderbilt Health, Department of Pathology, Microbiology, and Immunology
1161 21st Avenue South, C-3317 MCN, Nashville, TN 37232-2561
Telephone: 615-343-5708
E-mail: james.h.nichols@vumc.org

Keywords

Nicotine, point of care, toxicology

Abstract

Background: Tobacco use adversely affects surgical outcomes, prompting guidelines for preoperative smoking cessation. Accurate assessment of smoking status is critical, yet self-reporting is often unreliable. Biochemical verification via nicotine metabolites, particularly cotinine, offers objective evaluation. While mass spectrometry provides high sensitivity, point-of-care (POC) tests offer rapid, accessible alternatives. This study assesses the performance of NicCheck I, a CLIA-waived urine test, by comparing its results to liquid chromatography–tandem mass spectrometry (LC-MS/MS) for detecting tobacco exposure in surgical patients.

Methods: Twenty-four frozen urine samples (12 cotinine-negative, 12 cotinine-positive) were analyzed using NicCheck I test strips and LC-MS/MS. Sample integrity, and pH were assessed using Siemens Multistix 10 SG reagent strips.

Results: NicCheck I correctly identified all negative samples but detected nicotine metabolites in only 9 of 12 LC-MS/MS-confirmed positive samples. False negatives were associated with metabolite concentrations below the NicCheck I detection threshold (200 ng/mL) or urine pH outside the optimal range (4.5 – 8.5) consistent with manufacturer's package insert. LC-MS/MS demonstrated superior sensitivity.

Conclusion: NicCheck I test strips provide a practical and same-day rapid screening tool for nicotine exposure in clinical settings prior to surgery, but the test is less sensitive than LC-MS/MS. NicCheck I test strip performance may be affected by abnormal urine pH.

Introduction

Tobacco use remains a major public health challenge, and a leading preventable cause of disease, death, and disability both in the United States and globally. In the United States alone over 16 million people are living with a smoking-related disease resulting in more than 480,000 deaths each year [1].

In surgical care settings, smoking is strongly associated with adverse perioperative outcomes, including poor wound healing, higher risk of infection, cardiac and pulmonary complications, longer hospital stays, and increased overall healthcare costs [2,3,4]. Recognizing these risks, the World Health Organization recommends smoking cessation at least four weeks prior to surgery, noting that each additional week of abstinence reduces postoperative complications by approximately 19% [5]. While the optimal timeline for quitting smoking before surgery remains debated, the American Society of Anesthesiologists Task Force on Smoking Cessation emphasizes that patients should abstain from smoking for as long as possible both before and after surgery [6].

Accurate assessment of a patient's smoking status is essential for perioperative risk stratification and surgical planning. Self-reported smoking history, while commonly used, is often unreliable due to underreporting or misrepresentation, particularly in settings where smoking may influence surgical eligibility or outcomes [7, 8]. Biochemical verification through detection of nicotine and/or its metabolites in blood or urine offers a more objective and reliable alternative.

Nicotine, the primary alkaloid in tobacco, has a very short half-life (1-4 hours). In humans, approximately 70-80% of nicotine is rapidly metabolized by the cytochrome P450 2A6 (CYP2A6) in the liver to its primary metabolite, cotinine, which has a longer half-life (16-20 hours). Approximately 10-15% of the cotinine is excreted in the urine. Remaining cotinine is further metabolized to six primary metabolites including trans-3'-hydroxycotinine (3-OH-cotinine) which has a half-life of ~15 hours [9]. Cotinine and 3-OH-cotinine are considered reliable biomarkers for tobacco use due to their specificity, extended half-lives, and elevated urinary concentrations [10]. Importantly, cotinine concentrations in urine, saliva or serum are reflective of the amount of nicotine exposure [11].

Tobacco smoke exposure can be detected using various biochemical techniques, including gas chromatography (GC) and liquid chromatography (LC) coupled with mass spectrometry (MS), as well as immunological assays. Among these, MS-based methods offer high sensitivity and specificity for identifying nicotine and its metabolites. However, MS requires technical expertise, has long turnaround times, is expensive and is not widely available in many clinical laboratories. In contrast, point-of-care (POC) testing devices have emerged as practical alternatives for rapid, same-day screening for tobacco consumption. These Food and Drug Administration (FDA)-cleared and Clinical Laboratory Improvement Amendments (CLIA) -waived tests utilize colorimetric detection to provide qualitative or semi-quantitative results for nicotine metabolites in urine [8].

Given the importance of timely and accurate nicotine screening, particularly in plastic and reconstructive surgery where smoking cessation is often mandated, POC testing represents a valuable tool in perioperative care [12].The efficiency of NicCheck I (Mossman Associates, Inc, Milford, MA) test strips has previously been assessed against breath carbon monoxide (CO) and has since achieved a CLIA waived status [13]. This study aims to evaluate the analytical performance of NicCheck I test by comparing results for urine nicotine and its metabolites with those obtained through liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Methods

Specimens

Frozen urine samples (n=24) were obtained from ARUP laboratories (Salt Lake City, UT). These included negative (n=12) samples for cotinine by Enzyme Multiplied Immunoassay technique and positive (n=12) samples for nicotine and metabolites confirmed by quantitative LC-MS/MS.

Testing

Frozen urine samples were thawed at room temperature and centrifuged (2000 rpm for 5 minutes). NicCheck I positive/negative cotinine controls and urine samples were tested using the NicCheck I test strips following manufacturer's instructions in the package insert. The NicCheck I test strip detects nicotine in urine through a chemical reaction. Urine activates reagents on the strip, producing cyanogen chloride, which reacts with nicotine and/or its metabolites. This forms glutaconaldehyde, which then reacts with diethylthiobarbituric acid on the strip, resulting in a pink color that indicates the presence of nicotine and/or metabolites. Color developed on the test strip should be observed after 15 minutes and the intensity of color development can be compared to a color card provided by the manufacturer. Color intensity in the range of 1-6 or 7-12 as indicated on the color card are classified as low consumer and high consumer of nicotine respectively.

Sample pH and integrity were simultaneously assessed on Clinitek status analyzer using Siemens Multistix 10 SG reagent strips (Tarrytown, NY).

Results

NicCheck I test strips incubated with cotinine negative urine samples (n=12) did not develop any color after 15 minutes, indicating absence of nicotine and/or metabolites. NicCheck I strips incubated with LC-MS/MS confirmed positive urine samples developed a pink color in 9 out of the 12 samples indicating the presence of nicotine and/or metabolites (Table 1). According to the manufacturer's instruction for use, NicCheck I test strips can detect cotinine >200 ng/mL (as determined by Gas chromatography) with a positive predictive value of 98.4%.

Table 1: Association between qualitative color development on NicCheck I test strips, urine pH, and LC-MS/MS–measured cotinine and nicotine metabolite levels in urine samples (n=12) that tested positive for nicotine and its metabolites by LC-MS/MS.

Sample	Color development (range)	Urine pH (Clinitek)	Cotinine (ng/mL) (LC-MS/MS)	Nicotine and metabolites (ng/mL) (LC-MS/MS)
1	**Absent**	**6**	**32**	**<102**
2	Present (1)	5	330	988
3	Present (1)	5.5	587	1889
4	**Absent**	**6**	**17**	**<83**
5	Present (6-7)	6	1299	>4705
6	Present (1-2)	5.5	595	1560
7	Present (5-6)	5.5	1007	>4464
8	Present (7-8)	7	803	>3584
9	**Uneven coloration**	**≥9**	**1063**	**>3156**
10	Present (6-7)	5.5	1292	>3759
11	Present (2-3)	6.5	771	>2820
12	Present (1)	5	378	>2412

Samples 1 and 4 showed no color development and had low or near-below-detection analyte levels (<200 ng/mL), whereas increasing color intensity was associated with progressively higher cotinine and nicotine metabolite concentrations. Uneven coloration was observed in one urine sample (sample 9; pH ≥9) even with high analyte levels.

Discussion

With increasing awareness of the detrimental effects of smoking on postoperative outcomes, there is a growing need for rapid and simple POC testing to detect nicotine and its metabolites on the day of and immediately prior to surgery. Such testing has the potential to enhance perioperative decision-making, improve patient compliance, and ultimately optimize surgical outcomes. However, the availability of CLIA-waived tests for nicotine detection remains limited. To date, no studies have directly compared the performance of NicCheck I test strips with the LC-MS/MS.

In this study, we evaluated the analytical performance of NicCheck I test strips relative to LC-MS/MS. Our findings suggest that NicCheck I strips are adequate for initial screening of urine samples for nicotine and its metabolites. However, their sensitivity is inferior to LC-MS/MS, particularly for samples with lower concentrations of total nicotine metabolites which may result in false-negative results. This limitation underscores the importance of confirmatory testing in clinically ambiguous cases.

Additionally, urine pH was identified as a potential confounding factor, consistent with the manufacturer's package insert. One sample with a pH ≥9 yielded an inconclusive NicCheck I result despite having high concentrations of nicotine and its metabolites (>3156 ng/mL) as confirmed by LC-MS/MS. This observation suggests that urine pH should be routinely assessed when interpreting test results obtained using NicCheck I test strips.

Importantly, NicCheck I test strip does not distinguish between different tobacco sources or passive smoke exposure. For patients utilizing nicotine replacement therapy (NRT), confirmation of tobacco abstinence requires testing for anabasine, a tobacco-specific minor alkaloid absent in NRT products. LC-MS/MS remains the preferred method for anabasine detection and should be employed when precise differentiation is clinically necessary. Additionally, NicCheck I test strips detect all nicotine metabolites - including 3-hydroxycotinine which is the most abundant metabolite of nicotine (~40%) in urine and may persist for up to two weeks after smoking cessation [14]. This might cause some individuals to test positive even after recent abstention from smoking. Nonetheless, given that smoking cessation is typically recommended for at least 4 weeks if not more prior to surgery, the test retains clinical utility in identifying recent tobacco exposure [5,12].

One limitation of this evaluation was the small sample size. However, the number we evaluated was sufficient to delineate the differences in cutoff concentration between NicCheck I and LC-MS/MS methods, as well as define the potential for interference from urines with pH outside manufacturer recommended ranges. In conclusion, while NicCheck I test strips offer a reliable and rapid screening tool for nicotine exposure in the perioperative setting, their limitations necessitate cautious interpretation. When clinical decisions rely

on accurate smoking status - particularly in patients using NRT or presenting borderline results, LC-MS/MS should be utilized to ensure diagnostic precision.

Author Contributions

Seema Khattri Bhandari (Conceptualization- equal, investigation-lead, formal analysis-equal, data curation-equal, writing- original draft- lead). James H Nichols (Conceptualization-equal, formal analysis- equal, data curation-equal, Writing- review and editing, supervision-lead).

Conflicts of interest

The authors declare that they have no conflicts of interest related to this work.

Funding

The authors have no funding to declare pertaining to this article.

Data availability

The data supporting the findings of this article can be made available upon request.

References

1. Cigarette smoking, 2024. Available from https://www.cdc.gov/tobacco/about/index.html (accessed: 02/05/2025)
2. Fan Chiang YH, Lee YW, Lam F, Liao CC, Chang CC, Lin CS. Smoking increases the risk of postoperative wound complications: A propensity score-matched cohort study. Int Wound J. 2023;20: 391-402. https://doi.org/10.1111/iwj.13887
3. Tang E, Rodriguez RM, Srivastava A, Malhan R, Laksono I, Yan E et al. Impact of short duration smoking cessation on post-operative complications: A systematic review and meta-analysis. Journal of clinical anesthesia 2025; 106:111967. https://doi.org/10.1016/j.jclinane.2025.111967
4. Myers K, Hajek P, Hinds C, McRobbie H. Stopping Smoking Shortly Before Surgery and Postoperative Complications: A Systematic Review and Meta-analysis. Arch Intern Med 2011; 171:983–989. doi:10.1001/archinternmed.2011.97
5. Smoking greatly increases risk of complications after surgery, 2020. available from: https://www.who.int/news/item/20-01-2020-smoking-greatly-increases-risk-of-complications-after-surgery (accessed: 02/10/2025)
6. Statement on Smoking Cessation. Task force on smoking cessation, 2023. Available from https://www.asahq.org/standards-and-practice-parameters/statement-on-smoking-cessation.(accessed: 02/10/2025)
7. Alberg AJ, Worley ML, Tooze JA, Hatcher JL, Carpenter MJ, Day TA et al. The Validity of Self-reported Recent Smoking in Head and Neck Cancer Surgical Patients. Otolaryngol Head Neck Surg. 2015 Dec;153 :990-995. https://doi.org/10.1177/01945998155943
8. Benowitz, NL, Bernert JT, Foulds J, Hecht SS, Jacob P, Jarvis MJ, et al. Biochemical Verification of Tobacco Use and Abstinence: 2019 Update. Nicotine & tobacco research: official journal of the Society for Research on Nicotine and Tobacco, 2020;22: 1086–1097. doi:10.1093/ntr/ntz132.
9. Benowitz NL, Hukkanen J, Jacob P 3rd. Nicotine chemistry, metabolism, kinetics and biomarkers. Handb Exp Pharmacol. 2009; 192:29-60. https://doi.org/10.1007/978-3-540-69248-5_2
10. Gruzdys V, Signorelli H, Johnson-Davis KL. NicAlert™ test strip performance comparison with LC-MS/MS and immunoassay methods for nicotine and cotinine. Arch Clin Toxicol 2020;2:19-24. https://doi.org/10.46439/toxicology.2.010
11. Sharma P, Sane N, Anand SD, Marimutthu P, Benegal V. Assessment of cotinine in urine and saliva of smokers, passive smokers, and nonsmokers: Method validation using liquid chromatography and mass spectrometry. Indian J Psychiatry. 2019; 61:270-276. 10.4103/psychiatry.IndianJPsychiatry_61_18
12. Theocharidis V, Economopoulos KP. Smoking cessation prior to elective plastic surgery: why, when and how? Tob Induc Dis. 2014 ;12(Suppl 1): A18. doi: 10.1186/1617-9625-12-S1-A18
13. Leischow SJ, Merikle EP, Cook G, Newman R, Muramoto M. An evaluation of NicCheck I: a dipstick method for analyzing nicotine and its metabolites. Addict Behav. 1999;24(1):145-148. https://doi.org/10.1016/S0306-4603(98)00042-2
14. Benowitz NL, Jacob P 3rd. Trans-3'-hydroxycotinine: disposition kinetics, effects and plasma levels during cigarette smoking. Br J Clin Pharmacol. 2001;5: 53-59. doi: 10.1046/j.1365-2125.2001.01309.x

Case Report

Biochemical Clues in unravelling the Puzzle of Anti- synthetase Syndrome

Lekha Priyadharshini Kamarajan[1], Mala Mahto[2*], Sushil Kumar[2], Deependra Kumar Rai[3]

[1]*MBBS, PGT, Department of Biochemistry, AIIMS Patna, Patna, Bihar, India*
[2]*Biochemistry, AIIMS Patna, Patna, Bihar, India*
[3]*Department of Pulmonary Medicine, AIIMS Patna*

Article Info

Corresponding Author:
Mala Mahto
Biochemistry Department, AIIMS Patna Bihar 801507
ORCID ID:0000-0003-0445-6972
E-mail: dr.malamahto@gmail.com
Mobile No.:+91-9771404617

Keywords

Anti-synthetase syndrome, Myositis, Usual Interstitial pneumonia, Anti Jo-1

Abstract

Anti-synthetase syndrome (ASSD or ASyS) is a rare, systemic autoimmune disorder characterized by the presence of anti-aminoacyl transfer RNA synthetase (anti-ARS) antibodies, most commonly anti-Jo-1. It presents with a combination of myositis, interstitial lung disease (ILD), arthritis, Raynaud's phenomenon, mechanic's hands, and constitutional symptoms, often posing significant diagnostic challenges. We report the case of a 49-year-old male carpenter who presented with persistent cough, progressive dyspnoea, proximal muscle weakness, and joint pain. High-resolution computed tomography (HRCT) of the chest revealed interstitial lung disease with bilateral basal end-inspiratory crepitations noted on physical examination. Serological testing demonstrated anti-Jo-1 antibody positivity, confirming the diagnosis of anti-synthetase syndrome. The patient was initiated on glucocorticoids, hydroxychloroquine, and steroid-sparing immunosuppressive therapy, resulting in symptomatic and functional improvement. This case underscores the importance of early recognition of anti-synthetase syndrome, particularly in patients presenting with unexplained ILD and musculoskeletal symptoms. Delayed diagnosis is often attributed to limited clinical awareness and the restricted availability of routine testing for anti-ARS antibodies, particularly anti-Jo-1, which significantly affects patient outcomes. ILD is a predominant and serious manifestation of ASyS and remains a strong predictor of disease-related morbidity and mortality. Anti-synthetase syndrome should be considered in patients with interstitial lung disease, muscle weakness, and inflammatory arthritis, especially when anti-Jo-1 antibodies are detected. Early diagnosis and immunosuppressive therapy are critical to improving prognosis and preserving lung function.

Introduction

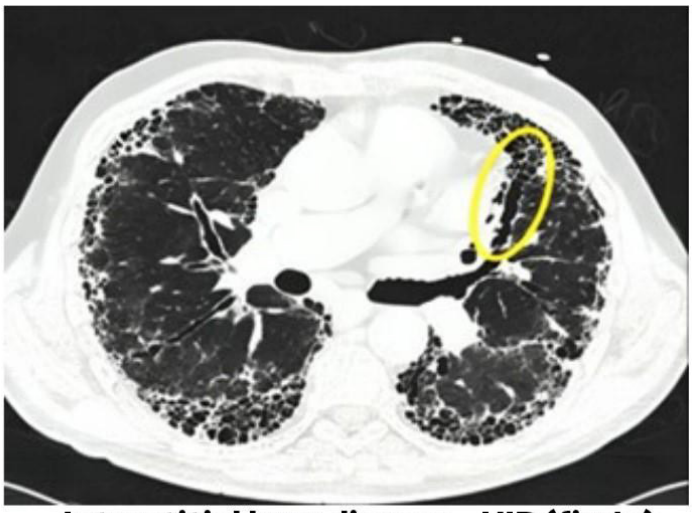
Anti-synthetase syndrome (ASS), also referred to as ASSD or ASyS, is a rare autoimmune disorder characterized by the presence of autoantibodies targeting aminoacyl-tRNA synthetases, which are essential enzymes in protein synthesis. The clinical spectrum of ASS/ASSD/ASyS typically includes proximal muscle weakness, polyarthritis, fever, Raynaud's phenomenon, interstitial lung disease (ILD), inflammatory myopathy, and hyperkeratotic changes on the fingertips [1,2]. Six specific anti-aminoacyl tRNA synthetase autoantibodies have been identified, including anti-Jo-1 (Histidyl), anti-PL7 (Threonyl), anti-PL12 (Alanyl), anti-OJ (Isoleucyl), anti-EJ (Glycyl), and anti-KS (Asparaginyl) [3]. These autoantibodies are not only central to defining the syndrome but also assist in correlating clinical phenotypes and underscore the significance of early diagnosis and individualized treatment approaches for effective ILD management [4].

Case Report

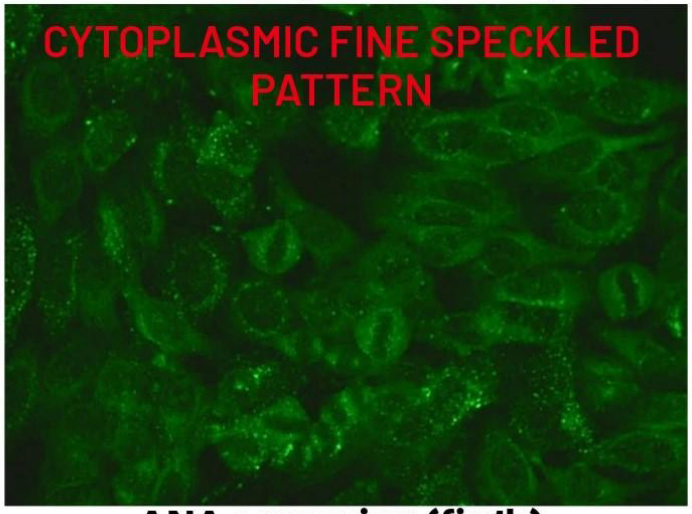
A 49-year-old male carpenter presented with a one-year history of persistent non-productive cough, progressive dyspnoea on exertion, generalized muscle weakness for four months, and multiple joint pains persisting for one year.

Clinical examination revealed synovitis of the metacarpophalangeal joints, proximal muscle weakness, and hyperkeratotic, fissured thickening along the radial aspects of the fingertips, characteristic of Mechanic's Hands. Auscultation of the chest revealed basal end-inspiratory crepitations, while vital signs remained within normal limits. Pulmonary function testing demonstrated a restrictive ventilatory defect with impaired gas transfer. High-resolution computed tomography (HRCT) of the chest confirmed interstitial lung disease (ILD) exhibiting an Usual Interstitial Pneumonia (UIP) pattern, characterized by honeycombing and traction bronchiectasis [Figure 1a]. Antinuclear antibody (ANA) screening was performed using indirect immunofluorescence (IIF) on human epithelial cells (HEp-2) and liver (Monkey) substrates (EUROIMMUN, Lübeck, Germany) at a 1:100 dilution. The fluorescence pattern observed was highly suggestive of anti-Jo-1 antibody positivity [Figure 1b]. An anti-synthetase syndrome autoantibody panel, including anti-KS, anti-EJ, anti-OJ, anti-Jo-1, anti-PL-7, and anti-PL-12, was performed using enzyme-linked immunosorbent assay (ELISA) (Bio-Rad, California, USA). The patient tested positive for both anti-Jo-1 (Histidyl) and anti-PL-12 (Alanyl) antibodies, while anti-PL-7, anti-EJ, and anti-OJ were negative. Laboratory findings demonstrated an elevated erythrocyte sedimentation rate, serum creatine kinase (CK) of 700 U/L (normal range: 40–320 U/L), and lactate dehydrogenase (LDH) of 500 U/L (normal range: 135–225 U/L), measured using AU series analysers (Beckman Coulter, USA). Based on clinical, radiological, and immunological findings, a diagnosis of Anti-Synthetase Syndrome (ASyS) was established. The patient was commenced on glucocorticoids and steroid-sparing immunosuppressants. Prednisolone was initiated at 50 mg daily and tapered by 10 mg every two weeks. Azathioprine was introduced at a dose of 100 mg per day (approximately 1.54 mg/kg/day) along with hydroxychloroquine 300 mg once daily. At the last follow-up after initiation of immunosuppressive therapy, the patient demonstrated significant clinical improvement with reduction in dyspnoea, joint pain, and proximal muscle weakness. Repeat laboratory evaluation also showed improvement in serum CK and LDH levels, supporting a favourable therapeutic response.

Figure 1 (A) and 1(B): Depicting CT scan finding revealing UIP pattern and ANA screening pattern on HEp-2 substrate reveals cytoplasmic pattern suggestive of anti Jo-1 antibodies.



Interstitial lung disease –UIP (fig 1a)



ANA screening (fig1b)

Discussion

Anti-Synthetase Syndrome (ASyS), also referred to as Anti-Synthetase Syndrome Disorder (ASSD), represents a rare and distinct subset of idiopathic inflammatory myopathies (IIMs), primarily defined by the presence of anti-aminoacyl tRNA synthetase (anti-ARS) autoantibodies along with a constellation of clinical features including interstitial lung disease (ILD), myositis, arthritis, Raynaud's phenomenon, mechanic's hands, and fever. Initially regarded as a myositis marker, the role of anti-ARS antibodies has significantly evolved over time to be recognized as a central diagnostic hallmark of ASyS, reflecting its multi-organ involvement and overlap with other connective tissue diseases [1]. Despite the growing understanding of ASyS, diagnosis and management remain complex due to the heterogeneity in clinical presentation, variability in antibody profiles, and overlapping symptoms with other autoimmune disorders. Anti-ARS antibodies are considered central to the pathogenesis and diagnosis of ASyS [2]. These insights explain the autoimmune nature of anti-ARS antibodies and their pathogenic potential in diseases like ASyS.

The patient's presentation with chronic cough, dyspnoea, arthritis, and muscle weakness prompted consideration of several differential diagnoses including rheumatoid arthritis-associated ILD, idiopathic pulmonary fibrosis, polymyositis/dermatomyositis, and systemic sclerosis-associated ILD. The UIP pattern on HRCT initially favoured fibrotic ILD, while inflammatory arthritis suggested a connective tissue disorder. However, elevated CK and LDH levels, proximal muscle weakness, and mechanic's hands pointed toward inflammatory myopathy-associated ILD. Confirmation was achieved through positive anti-Jo-1 and anti-PL-12 antibodies, establishing the diagnosis of Anti-Synthetase Syndrome.

Clinically, ASyS is notable for its diverse phenotypic spectrum. Solomon et al. highlighted cases where anti-OJ and anti-KS antibodies were predominantly associated with isolated ILD, while anti-ZO-positive patients displayed multi-systemic

involvement, including arthritis, myositis, and Raynaud's phenomenon [3]. This diversity reiterates the diagnostic challenge of ASyS, as some patients present solely with pulmonary manifestations while others experience widespread systemic disease. Our case, exhibiting anti-Jo1 and anti-PL12 co-positivity, contributes further evidence to this heterogeneity and underlines the need for clinicians to maintain a high index of suspicion even in atypical presentations. One of the most comprehensive studies on ASyS was conducted by Cavagna et al., who reported the findings from an international multicenter cohort involving 828 patients across ten countries. Their research demonstrated that anti-Jo1 antibodies were the most prevalent (72%), followed by anti-PL7, anti-PL12, anti-EJ and anti-OJ at varying frequencies [4]. Interestingly, our patient demonstrated dual positivity for anti-Jo-1 and anti-PL-12 antibodies, a rare serological finding in Anti-Synthetase Syndrome. The coexistence of both antibodies likely contributed to the prominent overlap of interstitial lung disease, inflammatory arthritis, and myositis observed in this case, highlighting the clinical heterogeneity of ASyS. ASyS patients commonly exhibit features of ILD, myositis, and arthritis, with ILD often dominating the clinical picture, particularly in anti-Jo1-positive cases. Studies have reported that up to 70% of patients with anti-Jo1 antibodies develop ILD, reinforcing its significance in disease monitoring and prognosis [5-8]. Additionally, arthralgia and arthritis in ASyS can mimic rheumatoid arthritis, especially in seronegative cases, which can mislead clinicians and delay the correct diagnosis [9].
The lung involvement in ASyS is not uniform. Although NSIP is the most frequently reported radiological pattern in Anti-Synthetase Syndrome, our patient demonstrated a UIP pattern characterized by honeycombing and traction bronchiectasis, highlighting the heterogeneity of pulmonary manifestations in ASyS [10]. Autoantibody detection has undergone significant technological advancement. Historically, anti-ARS antibodies were identified using immunoprecipitation, where they were shown to co-precipitate tRNA along with synthetases. Modern techniques now include Indirect Immunofluorescence (IIF), ELISA, Addressable Laser Bead Immunoassay (ALBIA), Fluorescent Enzyme Immunoassay (FEIA), and Chemiluminescent Immunoassay (CLIA) [11]. Among these, Line Immunoassays (LIA) have become a preferred diagnostic modality, allowing simultaneous detection of Myositis-Specific Antibodies (MSAs) and Myositis-Associated Antibodies (MAAs), enabling more efficient clinical characterization of IIMs. MSAs, including anti-Jo1, anti-Mi-2, anti-SRP, and anti-TIF1-γ, are typically specific to myositis subtypes, offering not only diagnostic but also prognostic insights. For instance, the detection of anti-Jo1 correlates strongly with ILD and guides the clinician towards early respiratory assessment. Laboratory findings significantly supported the diagnosis in our patient. Elevated CK and LDH levels indicated active muscle involvement, while the cytoplasmic ANA-IIF pattern provided an important clue toward anti-synthetase antibodies, particularly anti-Jo-1. Subsequent ELISA-based antibody profiling confirmed anti-Jo-1 and anti-PL-12 positivity. However, variations in sensitivity among different serological assays may affect antibody detection, especially in resource-limited settings. MAAs, such as anti-SSA/Ro, anti-PM-Scl, and anti-U1-RNP, while not unique to myositis, often indicate overlap syndromes with diseases like systemic lupus erythematosus (SLE) or systemic sclerosis [12-14]. ELISA-based myositis antibody profiling was useful in confirming anti-Jo-1 and anti-PL-12 positivity in our patient; however, assay-related limitations should be considered during interpretation. Variations in sensitivity and specificity among different serological platforms may influence antibody detection, and false-positive or cross-reactive results can occasionally occur, particularly in cases demonstrating dual antibody positivity. Therefore, serological findings should always be interpreted in correlation with clinical presentation, imaging findings, and other laboratory parameters. Despite the advances in serological testing, routine detection of anti-ARS antibodies apart from anti-Jo1 remains limited due to a lack of standardization, high costs, and absence of regulatory approval for many tests. This diagnostic gap is especially pronounced in resource-limited settings, reinforcing the importance of both clinical acumen and the need for improved, accessible diagnostic technologies in the future [11-14].
Additionally, Yehudina et al. described a compelling case of antisynthetase syndrome marked by ILD, Raynaud's phenomenon, digital ischemia, and severe occlusive vasculopathy. Their patient required a multi-modal treatment approach including corticosteroids, cyclophosphamide, intravenous immunoglobulin (IVIg) and rituximab to manage disease flares. Notably, the SARS-CoV-2 infection triggered ILD exacerbation, necessitating antifibrotic therapy with nintedanib, which resulted in improved lung function and clinical stability, underscoring the complex and relapsing-remitting nature of ASyS and the importance of individualized therapeutic strategies [15].

The patient described here fulfilled both the sets of criteria for the diagnosis of anti-synthetase syndrome proposed by Connors et al. and Solomon et al. Solomon et al [3], EULAR/ACR [16] and Connors et al [17] classification criteria have been included below Table 1.

Table 1: EULAR/ACR, Connors et al and Solomon et al classification criteria for Anti-synthetase syndrome.

EULAR/ACR 2017	Connors et al. (2010) Criteria	Solomon et al (2011) Criteria	Case Details
Definite IIM: score of at least 7.5 (8.7 with muscle biopsy) Probable IIM: score of at least 5.5 (6.7 with muscle biopsy)	Required: Presence of an anti-aminoacyl tRNA synthetase antibody	Required: Presence of an anti-aminoacyl tRNA synthetase antibody	Present
Anti-Jo-1 •Objective symmetrical weakness of upper and lower proximal extremities •Proximal leg muscles weaker than distal •Dysphagia or oesophageal motility disorders •Skin lesions (heliotrope rash, Gottron's papules, and Gottron's sign) •Age of onset •Typical features in muscle biopsy and elevated skeletal muscle enzymes	**PLUS, one or more of the following clinical features:** •Raynaud's phenomenon •Arthritis •Interstitial lung disease •Fever (not attributable to another cause) •Mechanic's hands (thickened and cracked skin on hands, particularly at fingertips)	**PLUS, two major or one major and two minor criteria: Major:** •Interstitial Lung Disease (not attributable to another cause) •Polymyositis or dermatomyositis by Bohan and Peter criteria **Minor:** •Arthritis •Raynaud's phenomenon •Mechanic's hands	**Findings:** •Interstitial Lung Disease (+) •Mechanic's hands (+) •Arthritis (+)

Anti-synthetase syndrome remains a diagnostically challenging entity due to its heterogeneous clinical and serological presentation. Our case highlights the rare coexistence of anti-Jo-1 and anti-PL-12 antibodies with UIP-pattern ILD, emphasizing the importance of integrating clinical, radiological, and serological findings for early diagnosis and management.

Conclusion

This case highlights the diagnostic challenge of Anti-Synthetase Syndrome in patients presenting with overlapping pulmonary and musculoskeletal manifestations. The rare coexistence of anti-Jo-1 and anti-PL-12 antibodies with UIP-pattern interstitial lung disease underscores the marked clinical and serological heterogeneity of ASyS. Recognition of elevated CK and LDH levels together with a characteristic cytoplasmic ANA-IIF pattern provided crucial biochemical clues that guided definitive serological diagnosis. Our case further emphasizes the importance of comprehensive antibody profiling and multidisciplinary evaluation for early diagnosis and timely immunosuppressive therapy, which are essential for improving clinical outcomes and preserving pulmonary function in ASyS. All authors have approved of the final draft for publication purpose and are accountable for the accuracy and integrity of the work.

Acknowledgements

We acknowledge the contributions of patient, their attendants, nuring officers and technical staff in providing us relevant kits to carry out the research.

Conflict of interest

None declared.

Funding

None.

Ethics approval

waived off for case reports at our institute.

Supplementary data

None.

Data availability statement

All relevant data have been provided with the manuscript.

Abbreviations

Anti-synthetase syndrome (ASSD or ASyS)
anti-aminoacyl transfer RNA synthetase (anti-ARS)
interstitial lung disease (ILD)
High-resolution computed tomography (HRCT)
Usual Interstitial Pneumonia (UIP)
Antinuclear antibody (ANA)
indirect immunofluorescence (IIF)
human epithelial cells (HEp-2)
enzyme-linked immunosorbent assay (ELISA)
creatine kinase (CK)
Lactate dehydrogenase(LDH)
idiopathic inflammatory myopathies (IIMs),
Addressable Laser Bead Immunoassay (ALBIA)
Fluorescent Enzyme Immunoassay (FEIA)
Chemiluminescent Immunoassay (CLIA)
Myositis-Specific Antibodies (MSAs)
Myositis-Associated Antibodies (MAAs)
Systemic lupus erythematosus (SLE)
Intravenous immunoglobulin (IVIg)
Connective tissue diseases (CTDs)

References

1. Mahler M, Miller FW, Fritzler MJ. Idiopathic Inflammatory Myopathies and the Anti-Synthetase Syndrome: A Comprehensive Review. Autoimmune Rev. 2014;13(0):367–371. Doi: 10.1016/j.autrev.2014.01.022.
2. Galindo-Feria, A. S., Notarnicola, A., Lundberg, I. E., & Horuluoglu, B. Aminoacyl-tRNA Synthetases: On Anti-Synthetase Syndrome and Beyond. Frontiers in Immunology, 13, 866087. Doi: 10.3389/fimmu.2022.866087.
3. Solomon, J., Swigris, J. J., & Brown, K. K. Myositis-related interstitial lung disease and antisynthetase syndrome. Jornal Brasileiro de Pneumologia, 37(1), 100–109. Doi: 10.1590/S1806-37132011000100015.
4. Bohan, A.; Peter, J.B. Polymyositis and dermatomyositis (first of two parts). N. Engl. J. Med. 1975, 292, 344–347
5. Cavagna L, Nuño L, Scirè CA, Govoni M, Lopez Longo FJ, Franceschini F, Neri R, et al. Clinical spectrum time course in anti Jo-1 positive anti-synthetase syndrome: results from an international retrospective multicenter study. Medicine. 2015;94(32) Doi: 10.1097/MD.0000000000001144.
6. Mosca M, Barsotti S, Neri R, Scirè C, Govoni M, Furini F, et al. Influence of Antisynthetase Antibodies Specificities on Antisynthetase Syndrome Clinical Spectrum Time Course. J Clin Med. 2019;8(11):2013. Doi:10.3390/jcm8112013.
7. Aggarwal R, Cassidy E, Fertig N, Koontz DC, Lucas M, Ascherman DP, et al. Patients with non-Jo-1 anti-tRNA-synthetase autoantibodies have worse survival than Jo-1 positive patients. Ann Rheum Dis. 2014; 73(1):227–232. Doi:10.1136/annrheumdis-2012-201800.
8. Patel HC, Lauder NN. The Anti-synthetase Syndrome. Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, Md. Am J Med. 2011;124(8):767–75. Doi: 10.1016/j.amjmed.2011.02.010.
9. Pinal-Fernandez I, Casal-Dominguez M, Huapaya JA, Albayda J, Paik JJ, Johnson C, et al. A longitudinal cohort study of the anti-synthetase syndrome: increased severity of interstitial lung disease in black patients and patients with anti-PL7 and anti-PL12 autoantibodies. Rheumatology. 2017;56(6):999-1007. Doi: 10.1093/rheumatology/kex021.
10. Ramos-Martinez E, Vega-Sánchez AE, Pérez-Rubio G, Mejia M, Buendía-Roldán I, González-Pérez MI, et al. Enhanced activity of NLRP3 inflammasome in the lung of patients with anti-synthetase syndrome. Cells. 2023;12(1):60. Doi: 10.3390/cells12010060.
11. Marco JL, Collins BF. Clinical manifestations and treatment of antisynthetase syndrome. Best. Pract. Res. Clin. Rheumatol. 2020, 34, 101503.
12. Marklund J, Hanna B, Jin T, Pullerits R. Myositis-associated antibodies predict the severity of lung involvement in adult patients with inflammatory myositis – a cohort study of 70 adult patients with myositis in a single center. Front Med. 2024;11. Doi: 10.3389/fmed.2024.1340310.
13. Cruellas MG, Viana VD, Levy-Neto M, de Souza FHC, Shinjo SK, et al. Myositis-specific and myositis-associated autoantibody profiles and their clinical associations in a large series of patients with polymyositis and dermatomyositis. Clinics. 2013;68(7):906–912. Doi: 10.6061/clinics/2013(07)04.
14. Nakashima R, Hosono Y, Mimori T. Clinical significance and new detection system of autoantibodies in myositis with interstitial lung disease. Lupus. 2016;25(9):925-933. Doi: 10.1177/0961203316645267.
15. Yehudina Y, Trypilka S, Isayeva A. Clinical puzzles and decision-making in antisynthetase syndrome. Cureus. 2021 Jun 25;13(6): e15931. Doi: 10.7759/cureus.15931.
16. Lundberg, I.E., Tjärnlund, A., Bottai, M., Werth, V.P., et al. 2017 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Adult and Juvenile Idiopathic Inflammatory Myopathies and Their Major Subgroups. *Arthritis Rheumatol.* 2017;69:2271–2282. https://doi.org/10.1002/art.40320.
17. Connors GR, Christopher-Stine L, Oddis CV, Danoff SK. Interstitial Lung Disease Associated With the Idiopathic Inflammatory Myopathies: What Progress Has Been Made in the Past 35 Years? *Chest*. 2015;148(6):1476-1487. Doi:10.1378/chest.10-018

Case Report

Infantile Osteopetrosis with a CLCN7 Variant: An Educational Case Highlighting Integrated Diagnosis and Variant Interpretation

Meena Shree A V[1], Ranjith Kumar Manokaran[2*], Dhaarani Jayaraman[3], Jai Prakash Srinivasan[3], Elayaraja Sivaprakasam[4]

[1]*Department of Pediatrics, Sri Ramachandra Institute of Higher Education and Research, Chennai*
[2]*Department of Pediatric Neurology, Sri Ramachandra Institute of Higher Education and Research, Chennai*
[3]*Department of Pediatric Hematology & Oncology, Sri Ramachandra Institute of Higher Education and Research, Chennai*
[3]*Department of Radiology, Sri Ramachandra Institute of Higher Education and Research, Chennai*
[4]*Associate Professor, Department of Pediatrics, Sri Ramachandra Institute of Higher Education and Research, Chennai*

Article Info

[*]***Corresponding Author:***
Ranjith Kumar Manokaran
Associate Professor,
Department of Pediatric Neurology,
Sri Ramachandra Institute of Higher Education and Research, Chennai, India
Telephone.: +91-9013860064
E-mail: doct.ranjith@yahoo.com

Keywords

Osteopetrosis, CLCN7 mutation, child, seizures, developmental delay

Abstract

Background: Infantile osteopetrosis is a rare inherited disorder caused by defective osteoclast-mediated bone resorption, resulting in increased bone density and marrow failure. The CLCN7 gene encodes a chloride channel essential for osteoclast acidification.

Case Presentation: A 4-month-old male presented with failure to thrive, pallor, hepatosplenomegaly, and recurrent infections. Examination revealed frontal bossing and hypotonia. Laboratory evaluation showed anemia and persistent hypocalcemia (6.8 mg/dL; reference 8.5–10.5 mg/dL). Skeletal survey demonstrated diffuse osteosclerosis with a characteristic “bone-within-bone” appearance. Differential diagnoses, including storage disorders and leukemia, were excluded. Clinical exome sequencing identified a homozygous CLCN7 missense variant, c.613G>A (p.Gly205Arg), classified as a variant of uncertain significance (VUS) based on ACMG/AMP criteria (PM2, PP3).

Conclusion: Diagnosis in this case was primarily clinico-radiological, with genetic findings providing supportive evidence. This report presents an educational case emphasizing the importance of integrated clinical, radiological, and laboratory evaluation, and highlights the challenges in interpreting variants of uncertain significance in clinical practice.

Background

Osteopetrosis comprises a group of inherited disorders characterized by defective osteoclastic bone resorption, resulting in increased bone density and impaired marrow function. Infantile autosomal recessive osteopetrosis represents the most severe form and typically presents early in life with anemia, hepatosplenomegaly, and skeletal abnormalities [1]. The CLCN7 gene, located on chromosome 16p13.3, encodes a chloride channel expressed in osteoclasts. This channel functions in conjunction with vacuolar H^+-ATPase to maintain an acidic microenvironment within the resorption lacuna, which is essential for bone mineral dissolution. Mutations in CLCN7 disrupt this process, leading to impaired bone resorption and progressive osteosclerosis [2]. This report is presented as an educational case highlighting the role of integrated diagnosis and the complexities of genetic variant interpretation

Case Presentation

A 4-month-old male infant presented with failure to thrive, pallor, hepatosplenomegaly, and recurrent infections. Birth history was unremarkable. On examination, the child had frontal bossing and generalized hypotonia (Figure 1 and 2).

Laboratory Evaluation

Complete blood count revealed anemia with preserved leukocyte counts. Serum calcium levels were persistently low (6.8 mg/dL; reference range: 8.5–10.5 mg/dL).
Further biochemical evaluation demonstrated elevated parathyroid hormone (PTH: 136 pg/mL), low-normal vitamin D levels (21 ng/mL), and mildly elevated alkaline phosphatase (ALP: 365 IU/L).

Table 1: Key Laboratory Findings and Interpretation in Infantile Osteopetrosis.

Parameter	Value	Reference Range	Interpretation
Hemoglobin	7.2 g/dL	10.5-13.5	Anemia from marrow encroachment
Albumin-corrected Ca	6.8 mg/dL	8.5-10.5	Hypocalcemia (impaired resorption)
PTH	136 pg/mL	15-65	Secondary hyperparathyroidism
25-OH Vitamin D	21 ng/mL	30-100	Low-normal (not primary cause)
ALP	365 IU/L	100-350	Elevated bone turnover
WBC	6660	Cells/ Cu mm	Rules out leukemia

The observed hypocalcemia, in conjunction with elevated PTH, is consistent with secondary hyperparathyroidism resulting from impaired osteoclastic bone resorption. Reduced calcium mobilization from bone leads to compensatory PTH elevation, while ALP elevation reflects increased bone turnover activity. These findings support the underlying pathophysiology of osteopetrosis.

Radiological Findings

Skeletal survey demonstrated diffuse osteosclerosis with classical metaphyseal "bone-within-bone" appearance and generalized increase in bone density, strongly suggestive of osteopetrosis.

Differential Diagnosis

Initial differentials included storage disorders (e.g., Gaucher disease) and hematological malignancy. These were excluded based on normal β-glucocerebrosidase enzyme assay, absence of abnormal cells on peripheral smear, and lack of organ-specific biochemical abnormalities, in conjunction with characteristic radiological findings.

Genetic Analysis

Clinical exome sequencing performed at MedGenome Laboratories identified a homozygous missense variant in CLCN7: c.613G>A (p.Gly205Arg) (transcript ENST00000382745.9).
The variant was absent (or extremely rare) in population databases such as gnomAD and has not been definitively reported in ClinVar. In silico prediction tools (SIFT, PolyPhen-2, MutationTaster) suggest a deleterious effect on protein function.
Based on ACMG/AMP criteria (PM2, PP3), the variant was classified as a variant of uncertain significance (VUS). Parental testing could not be performed due to logistical constraints, limiting segregation analysis.
Copy number variation analysis was not performed as part of this assay. The sequencing achieved standard clinical exome coverage with adequate read depth for variant detection in coding regions of CLCN7; however, limitations inherent to exome sequencing, including inability to detect deep intronic variants and structural rearrangements, must be considered.

Figure 1: Clinical image showing mild bilateral proptosis,microcephaly and pallor.

To
The Editor-in-Chief
The Electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine (eJIFCC)

Subject: Parental Consent for Use of Child's Photograph and clinical details in Manuscript Submission

Dear Editor,

I, the undersigned, am the parent/legal guardian of **Master Kavi Darshan**, who is featured in the manuscript entitled **"Infantile Osteopetrosis with a CLCN7 Variant: An Educational Case Highlighting Integrated Diagnosis and Variant Interpretation"** submitted for consideration for publication in the *Electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine (eJIFCC)*.

I hereby give my full consent for the inclusion and publication of my child's **photograph** and information in the above-mentioned article. I understand that the image may appear in print and/or electronic formats of the journal and that it may be accessible to the general public.

I confirm that:

- I have been informed about the purpose of including the photograph and understand that it will be used for academic and educational purposes.
- I have been assured that no identifying information other than the image itself will be disclosed.
- I understand that the image may be used in perpetuity as part of the published record and may not be withdrawn after publication.

I hereby authorize the *Electronic Journal of the International Federation of Clinical Chemistry and Laboratory Medicine* to use this photograph and clinical details in connection with the publication of the aforementioned article.

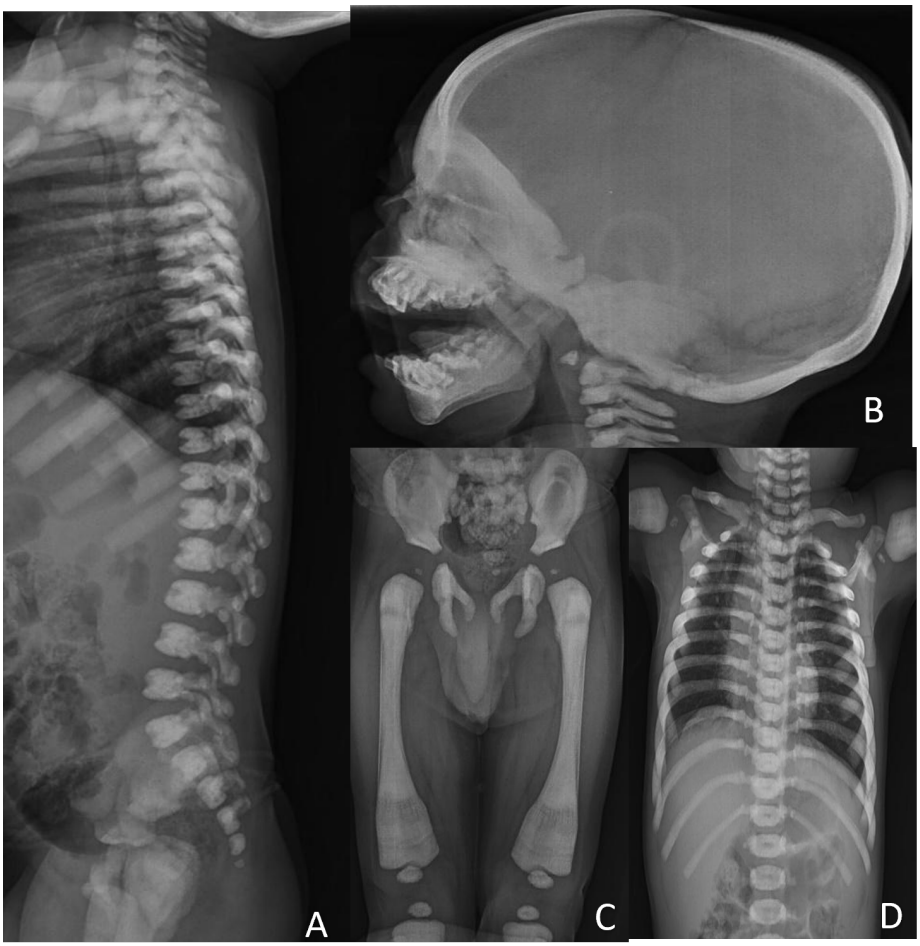
Thanking You,

Yours Sincerely,

Mr.Jagan S

(Father)

Figure 2: (A)DORSO LUMBAR SPINE: Homogenous increased density noted in vertebral bodies. (B) SKULL(LAT VIEW) :



Diffuse sclerosis involving the base of the skull,pituitary fossa,mastoid and cervical vertebra ; Atlanto axial widening present. (C)PELVIS : Squaring of ileum ; Flattened acetbulum ; Increased density of proximal femur ; Flaring of distal femur. (D) CHEST (PA VIEW) Diffuse increased density in the ribs, vertebral column and proximal humerus.

Discussion

Osteopetrosis is characterized by defective osteoclastic activity, leading to impaired bone resorption, skeletal sclerosis, and bone marrow failure. The diagnosis in this case was primarily established based on clinical features and characteristic radiological findings, with genetic analysis providing supportive evidence.

The CLCN7 gene plays a critical role in osteoclast function by facilitating chloride transport necessary for acidification of the resorption lacuna. Mutations disrupt this process, resulting in failure of bone mineral dissolution.[6]

Although the identified variant is classified as a VUS, the strong phenotypic concordance with CLCN7-related osteopetrosis supports a possible contributory role. However, the absence of segregation analysis and functional validation precludes definitive classification.

The variant was classified as a VUS based on ACMG/AMP criteria (PM2, PP3); however, the absence of stronger evidence limits definitive pathogenic classification. Additional criteria that could support reclassification include functional studies demonstrating impaired chloride channel activity (PS3), segregation analysis in affected family members (PP1), or identification of the same variant in unrelated affected individuals (PS4). In the absence of such data, cautious interpretation and clinicoradiological correlation remain essential.

The biochemical profile in this case further supports the diagnosis. Hypocalcemia in osteopetrosis results from defective osteoclastic activity, leading to reduced calcium release from bone. The elevated parathyroid hormone level observed reflects a compensatory response to hypocalcemia (secondary hyperparathyroidism). Mild elevation of alkaline phosphatase indicates ongoing bone remodeling activity, while vitamin D levels were within the low-normal range, excluding primary vitamin D deficiency as a cause of hypocalcemia. These laboratory findings aid in differentiating osteopetrosis from other conditions such as nutritional rickets, storage disorders, and hematological malignancies.[2,3]

From a laboratory medicine perspective, careful interpretation of biochemical parameters is essential. Serum calcium values were albumin-corrected; however, ionized calcium measurement, where available, provides a more accurate assessment of biologically active calcium. Pre-analytical variables such as sample handling and albumin levels may influence total calcium estimation. Analytical considerations, including assay methodology and inter-laboratory variability, should also be acknowledged

Biochemically, osteopetrosis must be differentiated from conditions such as nutritional rickets and hypoparathyroidism. Unlike rickets, where alkaline phosphatase is markedly elevated and hypocalcemia is often associated with elevated parathyroid hormone and low vitamin D, osteopetrosis typically demonstrates hypocalcemia with secondary hyperparathyroidism and only mild to moderate elevation of alkaline phosphatase. In hypoparathyroidism, hypocalcemia is associated with low or inappropriately normal parathyroid hormone levels, helping to distinguish it from osteopetrosis.

Besbas N et Al reported a 26-month-old child with osteopetrosis, neurodegeneration, retinal atrophy, and tubulopathy born to consanguineous parents who had CLCN7 gene with a novel homozygous R561Q variant. Both healthy parents were heterozygous for this amino acid substitution indicating autosomal recessive inheritance. The same homozygous nucleotide transition was found prenatally in a second child and the pregnancy was terminated at 17th week of gestation. A full autopsy was performed to the fetus, which confirmed the presence of osteopetrosis, thereby indicating that the variant observed indeed represents the disease-causing mutation.[5]

Hematopoietic stem cell transplantation (HSCT) remains the only curative treatment for severe infantile osteopetrosis. However, outcomes in CLCN7-related disease may be variable, particularly in cases associated with neurodegeneration, necessitating careful patient selection [4].

Follow up and Outcome

The child is currently on supportive management and remains clinically stable on follow-up. Hematopoietic stem cell transplantation (HSCT) was discussed with the family and is being considered as a potential therapeutic option. Regular monitoring for disease progression and complications is ongoing.

Clinical Implications

This case highlights the importance of maintaining a high index of suspicion for osteopetrosis in infants presenting with cytopenias, developmental delay, and characteristic skeletal changes.Early integration of clinical, laboratory, radiological, and genetic findings is essential for timely diagnosis and management. Identification of genetic variants should be interpreted cautiously, particularly when classified as VUS, and should be correlated with clinical findings.

Limitations

The identified CLCN7 variant is classified as a VUS, limiting definitive pathogenic interpretation. Lack of parental testing precluded segregation analysis, and no functional studies were performed to validate its biological impact. Additionally, limited database evidence and reliance on in silico predictions restrict robust genotype–phenotype correlation.

Take-home messages

Biochemical parameters (calcium, parathyroid hormone, and alkaline phosphatase) are essential for supporting the diagnosis of osteopetrosis and excluding mimicking conditions.

Variants of uncertain significance should be reported with supporting evidence from population databases and in silico prediction tools.

Timing and appropriateness of hematopoietic stem cell transplantation require careful consideration in CLCN7-related osteopetrosis

Author Contributions (CRediT Statement)

Meena Shree A V (MSAV): Conceptualization; Data Curation; Investigation; Writing – Original Draft; Visualization.

Ranjith Kumar Manokaran (RKM): Conceptualization; Methodology; Supervision; Writing – Review & Editing; Project Administration.

Dhaarani Jayaraman (DJ): Resources; Investigation; Writing – Review & Editing; Validation.

Jai Prakash Srinivasan (JPS): Resources; Visualization; Data Curation.

Elayaraja Sivaprakasam (ES): Clinical Management; Validation; Writing – Review & Editing.

All authors contributed to patient care, reviewed the manuscript critically for intellectual content, and approved the final version for submission.

The authors did not receive any financial support for the research, authorship, and/or publication of this article.

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Disclosures

None. The authors have no competing interests to declare that are relevant to the content of this article.

Conflicts of Interest

None.

Ethical statements

The study is in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Funding
The authors have no relevant financial or non-financial interests to disclose.

Author contributions
MSAV, RKM, DJ and JPS worked on the data analysis and wrote the initial draft; DJ, RKM and ES revised it for the clinical content and final revision for intellectual content by DJ and RKM. All the other authors were involved in the management of the child. All authors read and approved the final manuscript.

References
1. Stark Z, Savarirayan R. Osteopetrosis. Orphanet J Rare Dis. 2009;4:5.
2. Sobacchi C, Villa A, Schulz A, Kornak U. CLCN7-related osteopetrosis. In: Adam MP, Feldman J, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington; 1993–2024.
3. Del Fattore A, Cappariello A, Teti A. Genetics, pathogenesis and complications of osteopetrosis. Bone. 2008;42(1):19–29.
4. Schulz A, Moshous D. Hematopoietic stem cell transplantation in osteopetrosis. Bone. 2023;169:116645.
5. Besbas N, Draaken M, Ludwig M, et al. A novel CLCN7 mutation resulting in a most severe form of autosomal recessive osteopetrosis. Eur J Pediatr. 2009;168(12):1449–1454.
6. Kornak U, Kasper D, Bösl MR, et al. Loss of the ClC-7 chloride channel leads to osteopetrosis in mice and man. Cell. 2001;104(2):205–215

Case Report

The Truth Lies Within the Gap

Shaun Tan[1*], Alex Yong[1], Sharon Saw[1], Sunil Sethi[1]

[1]*Department of Laboratory Medicine, National University Hospital, Singapore*

Article Info

[*]***Corresponding Author:***
Shaun Tan
5 Lower Kent Ridge Road, MB Level 3, S119074
E-mail: shaun_sy_tan@nuhs.edu.sg

Keywords

Osmolality, Osmolal gap, Hyponatremia, Syndrome of inappropriate antidiuretic hormone secretion (SIADH)

Abstract

We report a case in which a post-analytical transcription error in serum osmolality led to diagnostic delay and unnecessary investigations in a patient with hyponatremia. The principal learning point of this case is that discrepancy between measured and calculated osmolality—the osmolal gap—was critical to detecting the error and redirecting the diagnostic evaluation. A 68-year-old female with a past medical history of hyperlipidemia, hypertension and type 2 diabetes mellitus, was scheduled for inpatient open reduction and internal fixation of a right distal radius fracture under the Hand surgery service. She was referred to the endocrinologists in view of persistent hyponatremia. Pre-operatively, her serum sodium was 128mmol/L (reference interval 135-145mmol/L), potassium was 4.2mmol/L (3.5-5mmol/L), urea was 4.8mmol/L (2.0-6.5mmol/L), random glucose was 5.2mmol/L (4.0-7.8mmol/L) and measured osmolality was 296 mOsm/Kg (275-300 mOsm/Kg). Records of her previous serum sodium trend were reviewed, and all were within the normal reference interval since 2011. In this case, the serum osmolality was within our laboratory's normal reference interval albeit at the higher end at 296 mOsm/Kg. The initial clinical impression at this point of time was thus normo-osmolar pseudohyponatremia. The clinical pathologist's opinion was requested, and the patient's serum sodium was further evaluated on the RAPIDLab® 1200 Systems Siemens blood gas analyser. This returned a similar sodium result of 128 mmol/L, excluding the possibility of electrolyte exclusion effect on the initial indirect ISE (Beckman Coulter, CA, USA). Upon further investigation, it showed that the osmolal gap was spuriously elevated (30mOsm/kg) due to a transcription error for the measured osmolality, hence confounding the true aetiology of hyponatremia. This case has shown that prompt communication between laboratory professionals and physicians will assist in resolving clinically discordant results. The osmolal gap, in particular, served as a useful adjunct in working up an ambiguous cause of hyponatremia.

Osmolality is an important laboratory marker reflecting the colligative property of a solution, measuring the solute concentration. It may be defined as the number of osmoles of solute per kilogram of solvent, and indicates the amount of solute particles dissolved in solution [1]. The osmolal gap (OG) is the difference between the measured osmolality (MO) and calculated osmolality (CO) and an increased gap suggests the presence of unmeasured, osmotically active particles. Where the MO is usually assessed by freezing point depression method on osmometers in most clinical chemistry laboratories, CO is determined using a number of osmolality equations [2]. A widely accepted and simple equation for CO initially proposed by Smithline et al, and used in our laboratory is: 2 x Sodium + Glucose + Urea (all in mmol/L) [3]. Hyponatremia remains the most common electrolyte disorder, seen in up to 30% of hospitalized individuals [4]. There are a multitude of potential aetiologies of hyponatremia, requiring meticulous work up by the physician and investigations often involve the determination of serum and urine osmolality as well as urine sodium.

The laboratory plays and important role in ensuring reliable results are reported and therefore interpreted correctly in such cases. Herein this report presents the evaluation of a surgical patient with hyponatremia posing a diagnostic conundrum to the clinical team, requiring a multi-disciplinary team approach with laboratory consult for successful resolution. Discussion of this case involves integrating clinical aspects of chemical pathology in the investigation of hyponatremia, osmolality concepts, as well as analytical understanding of the osmometer instrument.

Case Report

Case Description

A 68-year-old female with a past medical history of hyperlipidemia, hypertension and type 2 diabetes mellitus, was scheduled for open reduction and internal fixation of a right distal radius fracture. She was referred to the endocrinologists in view of persistent hyponatremia. Clinically, the patient was asymptomatic and did not report nausea, vomiting, giddiness, lethargy or any changes in her mental status. She indicated a recent increase in fluid intake from 500-800ml/day to 1000ml/day. At the time of admission, she was on Simvastatin 20mg, Lisinopril 10mg, Metformin 250mg. Pre-operatively, her serum sodium was 128mmol/L (reference interval 135-145mmol/L), potassium was 4.2mmol/L (3.5-5mmol/L), andom glucose was 5.2mmol/L (4.0-7.8mmol/L) and measured osmolality was 296 mOsm/Kg (275-300 mOsm/Kg). Records of her previous serum sodium trend were reviewed and all were within the normal reference interval since 2011. The patient demonstrated normal renal function with a serum creatinine of 66 µmol/L (reference interval 50–90 µmol/L) and urea of 4.8 mmol/L, indicating preserved renal function. These findings support the exclusion of renal insufficiency as a cause of hyponatremia.Physical examination revealed that she was clinically euvolemic with normal skin turgor, appearance of a slightly dry tongue but no pedal oedema or features suggestive of Cushing syndrome.

When approaching the problem of hyponatremia, it is helpful to analyse serum osmolality to determine the underlying cause. In this case, the serum osmolality was within our laboratory's normal reference interval albeit at the higher end at 296 mOsm/Kg. The initial clinical impression at this point of time was thus normo-osmolar pseudohyponatremia possibly contributed by use of TCMs. The clinical team had instructed the patient to cease taking her TCMs, but the hyponatremia persisted. Further investigations revealed total cholesterol of 3.93mmol/L, triglyceride level 0.67 mmol/L and LDL-cholesterol 2.24mmol/L. A myeloma panel performed did not detect abnormal bands on serum protein electrophoresis, with a normal serum protein level of 72 g/L (65-82) including that of albumin and globulin. The laboratory's opinion was requested, and the patient's serum sodium was further evaluated on the RAPIDLab® 1200 Systems Siemens blood gas analyzer, which is a direct Ion selective electrode (ISE) method. This returned a similar sodium result of 128 mmol/L, excluding the possibility of electrolyte exclusion effect on the initial indirect ISE (Beckman Coulter, CA, USA).

Additional investigations were ordered for endocrine evaluation of hyponatremia including a thyroid function test (TFT) and 250mcg short synacthen test. TFT demonstrated normal free thyroxine of 15.2 pmol/L (8-16pmol/L), and TSH of 0.71 mIU/L (0.45-4.5mIU/L). An appropriate cortisol response to the short synacthen test with a peak cortisol of 609nmol/L one-hour post-administration excluded adrenal insufficiency.

Laboratory Troubleshooting and Intervention

The "normal" measured serum osmolality results now appeared to be a red herring. Pseudohyponatremia was unlikely in view of the investigations performed above. Our laboratory team first proceeded to calculate the osmolality, as the clinical picture and patient's laboratory results appeared to be discordant. The results were carefully scrutinized and the calculated osmolality, based on the formula of 2 x Na + Glucose + Urea (all in mmol/L) indicated an elevated osmolal gap of 30 mOsm/kg (normal <10mOsm/Kg). The increased osmolal gap was immediately alerted to the endocrinologists but there was no clinical evidence to suggest any ingestion of osmotically substance, such as ethanol. This led to the suspicion of possible error in the initial reported serum osmolality of 269mOsm/kg. In our laboratory, a bench top Advanced Micro-Osmometer Model 3320 is used for osmolality measurement in urine and serum. It utilizes the technique of freezing-point depression to measure the osmolality of the solutions. A volume of 20uL of sample is inserted into the osmometer through a sampler with a new sampler tip. After one minute, the result will be displayed, recorded onto an osmometer worksheet and manually typed

onto the laboratory information system (LIS). The workflow for our osmolality samples, including the sample in question, is that they are analyzed only once before reporting by the medical technologist on duty. The initial sample earlier in the day was retrieved and reanalyzed three times for osmolality where the repeated results were 269 mOsm/Kg, consistent across three reruns. The initial result (296 mOsm/kg) was thus amended to 269 mOsm/kg.

Core Teaching Point: Post-analytical Transcription error

The large discrepancy between the initial result and the corrected reruns pointed to a post-analytical transcription error as the likely cause. Such errors typically occur when laboratory results are inaccurately transcribed from the worksheet into the LIS, leading to incorrect reporting. The diagnosis was thus changed from normo-osmolar hyponatremia to hypo-osmolar hyponatremia, which in a setting of post-fracture pain, was consistent with the syndrome of inappropriate anti-diuretic hormone secretion (SIADH). This was further supported by a urine sodium of 60mmol/L as well as inappropriately elevated urine osmolality (322mOsm/kg). The patient's fluid intake was restricted to 1.5L/day and she was followed up in outpatient clinic, with her serum sodium subsequently returning to normal. The table below summarizes the patient's key investigations and results:

Table 1: Summary of patient's clinical feature, investigations and laboratory troubleshooting.

Clinical features	Basic investigations	Endocrine function testing	Lab troubleshooting
•Distal radius fracture •Euvolemic hyponatremia with no changes in neurological state •No weakness, fatigue, confusion or drop in Glasgow coma scale	•Na 128 mmol/L (135-145) •Serum osmolality 296 mmol/kg (275-300) •K 4.2 mmol/L (3.5-5) •Urea 4.8 mmol/L (2-6.5) •Creatinine 66 umol/L (50-90) •Total Cholesterol 3.93mmol/L; Triglyceride 0.67mmol/L; HDL-C 1 mmol/L; LDL-C 2.24 mmol/L •Glucose 5.2mmol/L	•8am serum cortisol 285 nmol/L (123-626) •Peak cortisol of 609nmol/L post short synacthen test •fT4 15.2 pmol/L (8-16pmol/L), and TSH 0.71 mIU/L (0.45-4.5mIU/L)	Direct ISE on blood gas analyzer •Similar result to indirect ISE, Na 128mmol/L • Unlikely pseudohyponatremia Calculation of osmolal gap •Elevated osmolal gap of 30 mOsm/kg • Informed endocrinologists to exclude potential causes of elevated osmolal gap Retrieved initial serum and re-analyzed serum osmolality x 3 • Likely true result of 269mOsm/kg transcribed wrongly as 296mOsm/kg Quality improvement measures instituted and will be explained in discussion

Discussion

Approach to Hyponatremia

Cumming et al have described the prevalence of hyponatremia in their cohort of elderly fracture patients to be 14.2% [5]. The aetiology of hyponatremia was found to be multifactorial in 72.7% of patients and the commonest factors include thiazide diuretics, dehydration, proton pump inhibitors, SIADH and mirtazapine. SIADH will be elaborated in further detail below. Biochemically, hyponatremia may be defined as a decreased plasma sodium concentration. In our laboratory, the lower limit of the reference interval is 135mmol/L, in line with recommendations from Biswas et al [6].

Hyponatremia is classified as hypo-osmolar, hyper-osmolar, normo-osmolar and hence the most important initial step is the determination of plasma osmolality [7]. Normo-osmolar hyponatremia in the presence of normal glucose and urea should alert one to the high likelihood of pseudohyponatremia attributed by the electrolyte exclusion effect. This occurs as a consequence of excess lipid molecules or proteins, which occupy the solid fraction of serum, leading to interference [8]. The causes an underestimation of sodium in indirect ISE analysers, as there is a predilution step involved and the indirect ISE assumes that the solid fraction of serum is 7%. Performing sodium analyses on a direct ISE analyzer will therefore negate the effect of pseudohyponatremia. In this case, a myeloma screen, a lipid panel as well as a direct ISE analysis of sodium excluded the possibility of pseudohyponatremia.

Hyper-osmolar hyponatremia exists when there are increased quantities of other solutes in the extracellular fluid. This leads to extracellular shifts of water, or intracellular shifts of sodium to maintain the osmotic balance between the extracellular fluid and intracellular fluid. Common causes include hyperglycemia, mannitol or glycine.

Hypo-osmolar hyponatremia is a "true hyponatremia" in a sense that there is excess loss of sodium (depletional hyponatremia) or an increased extracellular fluid volume (dilutional hyponatremia). Patients with depletional hyponatremia are often hypovolemic and exhibit clinical signs such as hypotension, tachycardia and reduced skin turgor. Severe mental confusion with seizures may occur at values less than 105mmol/L [7]. An important further investigation in this

scenario would be urinary sodium. If urinary sodium is low (<10mmol/L), this indicates proper renal function and that the source of sodium loss would likely be from the gastrointestinal tract or skin. Urinary sodium losses of >20mmol/L generally suggest that renal loss is more likely. Dilutional hyponatremia occurs as a consequence of excess water retention and typically manifests as oedema on physical examination. Causes of dilutional hyponatremia include renal failure, congestive heart failure, cirrhosis as well as nephrotic syndrome. Hypo-osmolar hyponatremia in the context of euvolemic volume status in the patient occurs in SIADH, primary polydipsia and endocrine disorders such as adrenal insufficiency and hypothyroidism. An algorithm being used in our laboratory is provided in Figure 1.

Clinical aspects of SIADH

SIADH is the commonest cause of euvolemic hyponatremia in hospitalized patients [10]. It occurs as a result of unsuppressed secretion of the antidiuretic hormone (ADH, arginine vasopressin or AVP) resulting in inappropriately concentrated urine and therefore a reduced urine volume [11]. In most patients with SIADH, ingestion of water fails to adequately suppress ADH hence urine remains concentrated. This causes inappropriate and excessive free water retention, increasing total body water and subsequently dilutional hyponatremia. Subsequently the phenomenon is compounded by the body's response to water retention by suppressing renin-angiotensin-aldosterone axis, thus resulting in even greater renal sodium wasting [10]. As a result, a high urinary sodium excretion will be observed. The causes of SIADH can generally be divided into five main causes as described by Hannon et al in the European Journal of Endocrinology [12]. These include (1) Malignancies such as small cell lung cancer, sarcoma, lymphoma, gastrointestinal tract and pancreatic tumors; (2) Drugs commonly implicated are selective serotonin reuptake inhibitors, carbamazepine and tricyclic antidepressants; (3) Pulmonary causes such as Pneumonia, tuberculosis and vasculitis; (4) Intracranial pathology for example meningitis, gliomas, abscesses, and subarachnoid haemorrhage; (5) Miscellaneous causes for example post-operative pain, acute intermittent porphyria and Guillain-Barre syndrome.

For this case, in the absence of any drug aetiology, liver/renal impairment, or endocrinopathies such as thyroid/adrenal dysfunction, our team attributed her hyponatremia SIADH to being pain induced. The pathogenesis of pain leading to SIADH has been previously described by Mavani et al [13]. Arginine vasopressin, or ADH, is reported to have significant effects on pain perception in the central nervous system via the receptors V1a and V1b located within the neuronal cells and hypothalamus. In response to painful stimuli, central and peripheral ADH levels are reportedly increased to act as a potent analgesia, thereby increasing the body's pain threshold and reducing the nociceptive signals [14]. Post-fracture pain was identified as the precipitating cause of SIADH in this patient, in keeping with the known neurogenic stimulus for ADH release.

Biochemical and Clinical Diagnosis of SIADH

Based on the clinical guidelines from the European Society of Endocrinology, essential biochemical criteria for SIADH include low serum osmolality <275mOsm/kg, Urine sodium >30mmol/L with normal dietary salt intake and inappropriately elevated urine osmolality >100mOsm/kg. In addition, there the patient should be clinically euvolemic with the exclusion of adrenal, thyroid, or renal insufficiency with no recent use of diuretic agents [15]. These are summarized in table 2 below. Management SIADH includes initial measures such as fluid restriction and hypertonic saline. Drug treatments may be rendered for more severe SIADH including demecleocycline, an ADH inhibitor and Vaptans, which are ADH receptor antagonists. Earlier identification of the transcription error would have avoided unnecessary investigations, shortened the diagnostic evaluation, and enabled earlier initiation of targeted management with fluid restriction and pain control. In this case, our patient responded well to fluid restriction, pain control of the fracture, and close monitoring of serum sodium.

Table 2: Criteria for Diagnosis of SIADH.

Clinical Criteria	Biochemical Criteria	Exclusion cases
• Clinically euvolemic – No clinical signs of tachycardia, decreased skin turgor or dry mucous membranes – No oedema or ascites • No recent use of diuretics	• Serum Na <135mmol/L • Low serum osmolality <275 mOsm/kg • Urine Na >30mmol/L • Urine osmolality > 100 mOsm/kg	• No adrenal insufficiency • No hypothyroidism • No renal insufficiency

Clinical consequence – delay in care and unnecessary investigations

Of note, it is important to mention that our patient fulfilled the clinical criteria of SIADH but her diagnosis was confounded by an erroneous laboratory result. This had led to unnecessary further investigations and delay to timely diagnosis and appropriate management. For example, the normoosmolar hyponatremia led the clinical team down the rabbit hole of pseudohyponatremia which included work-up for lipids, myeloma screen and direct ISE testing. Thankfully there were no adverse outcomes and interventions in her case, but this underscores the fact that an estimated 70 percent of medical decision-making predicates on reliable laboratory test results [16]. As laboratory professionals, any discordant results with the clinical picture mandates us to do due diligence to thoroughly investigate the laboratory processes involved. In this case, our laboratory professionals picked up the discrepancy between measured and calculated osmolality in a timely fashion.

Osmolality – The Truth lies in the Gap and the discovery of Transcription error

When screening patients for the presence of exogenous osmotically active substances such as ethylene glycol and ethanol, the presence of an increased osmolal gap is important. This is calculated by subtracting calculated osmolality from measured osmolality. Choy et al have documented up to 37 different formulae for calculated osmolality, but detailed comparisons are beyond the scope of this report [1]. Nonetheless, a common and widely accepted formula adoped by my laboratory remains the Smithline-Gardner formula which uses 2 x Sodium + Glucose + Urea (All values in mmol/L) [3]. A normal osmolal gap should typically be less than 10 mOsm/Kg and an elevated gap indicates presence of unmeasured, osmotically active molecules such as ethanol, mannitol or hyperlipidemia. Calculated osmolality and osmolal gaps are not reported in our laboratory at present. Importantly, the key diagnostic clue in this case was the unexpectedly elevated osmolal gap, which prompted re-evaluation of the measured osmolality. Recognition of this discrepancy between measured and calculated osmolality triggered repeat analysis of the stored serum sample and ultimately led to the identification of a post-analytical transcription error. This case therefore highlights the value of the osmolal gap not only as a screening tool for unmeasured osmoles, but also as a practical internal consistency check for detecting laboratory errors when results appear discordant with the clinical picture. As described above, the elevated osmolal gap of 30 mOsm/kg prompted the laboratory to relay this information to the clinicians to first exclude intoxicating substances. Upon confirmation that this was clinically unlikely, we then proceeded to retrieve the initial serum from our refrigerated stockyard for analyses as we now suspected a wrong initial result. Based on a study by Bezuidenhout et al [17], it was concluded that serum and plasma osmolality were stable for up to 36 hours. The Advanced micro-osmometer utilizing the freezing-point depression method had been shown by Koumantakis to have excellent analytical precision (within-run Coefficient of variation of 0.59%, between-day Coefficient of variation of 0.58%). Moreover, based on our external quality assessment subscription, internal quality control processes and instrument maintenance charts there was no issue with our instrument. Reanalysis of the serum revealed that in all likelihood, there was a post-analytical transcription error of 296 mOsm/kg instead of 269 mOsm/kg which led to the result being retracted and amended. Data from Plebani's group has indicated that errors in the post-analytic phase, including that of transcription errors, have a relative frequency of 12.5-20% within the total testing process [18]. These were especially prevalent in scenarios where manual keyboard entry resulted in error, or analytical instruments that were not adequately interfaced with laboratory information systems and this was an apt description of our case in question.

Key Incident Monitoring and Management System (KIMMS)

More recently, Badrick et al have reported valuable data from the Australasian Quality improvement programme known as the Key Incident Monitoring and Management System (KIMMS) [19]. KIMMS records incidents known as process defects as well as episodes to calculate incident rates for pre- and post-analytical errors, thereafter using a failure mode effects analysis (FMEA) to assign a quantifiable risk and harm to the incident. A relevant category to our case was "Result retraction or correction" which described the occasion when result(s) have been modified because of an error after release in any form by the laboratory, and included data entry errors. Under KIMMS, result retraction or correction events are recognised as high risk–harm incidents, as they involve the potential dissemination of incorrect diagnostic information that may influence clinical decision-making, patient management, or treatment pathways. Among the top ten categories, KIMMS data from 2008-2016 showed that "result retraction or correction" was the tenth highest with 91,895 (3.09%) of incidents. However, its weighted risk-harm metric was the second highest ranking at 4,502,855 (12.01%). This indicates that a retracted laboratory report based on post-analytical error confers a high patient safety threat, which may result in further unnecessary investigations or treatments to a wrong diagnosis.

Allowable limits of difference and duplicate testing

As osmolality is tightly controlled by the body, it would be challenging to apply the commonly used analytical goal formula of $CVa < 0.5*CVi$ because the biological variability of osmolality reported in literature ranges from only 1.0-1.3% [20]. One option to mitigate total error would be to apply allowable limits of performance, which will allow my laboratory to assess its performance based on consensus

analytical performance goals. An available external quality assurance (EQA) programme, which my laboratory also subscribes to, would be the RCPAQAP General Serum Chemistry program. This has published extensive biochemistry data including that of osmolality. As outlined in the article by Jones et al [21]. such supporting information will allow users to determine if assays are meeting the requisite quality standard to diagnose or monitor patients within or between laboratories. For the use of serum or urine osmolality in diagnosis, the basis of limits given by RCPAQAP are that of minimal total error (bias and precision), in particular +/- 8 up to 266mOsm/L or 3% if more than 266 mOsm/L. Repeat testing, such as the one proposed by Deetz et al, involves automatically repeating laboratory tests when values exceed a critical threshold or trigger other automated rules such as a delta check in order to improve the accuracy and reliability of results [22]. A viable strategy that could be proposed would be to support osmolality analyses with duplicate testing to ensure the allowable limits of performance are not exceeded to mitigate error, and to provide warning of clinically important changes in assay performance. If these analytical performance limits were exceeded, it would then prompt investigation into one's quality controls and reporting processes and trigger repeating testing in order to troubleshoot potential error. If this had been performed for our case, we would have been able to pick up a discordant osmolality of an erroneous 296 mOsm/L versus a correct 269 mOsm/L result (difference of 27Mosm/L or 10% difference) and this would have saved the patient from unnecessary further investigations.

This case highlights several important considerations and learning points. If the result of an osmolal gap were readily visible to the clinicians on the electronic medical records, they could have significantly reduced the time taken to diagnose this patient. This could also potentially eliminate the need to order other redundant tests, resulting in cost savings. As such, our laboratory is considering the implementation of reportable calculated osmolality and osmolal gaps for all patients using programmable formulae on the laboratory information system (LIS).

Corrective and Preventive measures

We reviewed our current laboratory practice for medical technologists reporting osmolality results. Two medical technologists on duty are designated to perform osmolality testing. The first technologist performs the direct analyses and writes the result on an osmolality worksheet, whilst the second technologist verifies the results and transcribes the results onto the LIS. We suspect that the first technologist had transcribed the result wrongly onto the worksheet and the second technologist simply validated this result instead of checking on the osmometer screen. We improved our laboratory processes in order to reduce post analytical transcription error rate and enhance patient safety. The second technologist now has to verify the result directly displayed on the osmometer and on the osmolality worksheet before validating the result. Staff also were reminded and went through retraining to ensure competency in transcription of results into the LIS.

There was a need to review whether this was the appropriate osmometer model for our laboratory. Although our current purchased osmometer has not reached the end of its life-cycle, this case provided impetus for considering other models which would better suit our needs. The daily samples for urine/serum osmolality received by our laboratory remains more than 10 samples a day, with both adult and paediatric patients requiring osmolality determination. To reduce manual intervention, a consideration would be an osmometer from Advanced Instruments with integrated barcode scanner, automated liquid handling as well as middleware integrated with our LIS, known as A20 Advanced Automated osmometer [23]. This will be proposed during our next procurement process.

The summarized steps taken to mitigate and prevent further such occurences

1. Dual medical technologist verification at osmometer before transcription into the LIS
2. Retraining of staff to ensure competency in manual processes such as transcription for all analysers which are not interfaced
3. Proposal for automated osmolal gap calculation
4. Reviewing Osmometer instruments during our next procurement to include models which have the option of interfacing with the LIS

Conclusion

Taken together, this case has shown that prompt communication between laboratory professionals and physicians will assist in resolving clinically discordant results. It further emphasizes the importance of carefully examining pre-analytical, analytical and post-analytical laboratory processes for abnormalities. The osmolal gap, in particular, served as a useful adjunct in working up an ambiguous cause of hyponatremia. Reliable laboratory results must not be taken for granted, and chemical pathologists maintain an important role in bridging clinical details with technical laboratory aspects.

Declaration of Conflict of Interests and Disclosures

All authors declare that they do not have any conflicts of interest and nothing to disclose.

Ethical Approval

This incident report constituted part of our Laboratory's continuous quality improvement where errors and root cause analyses are performed. Ethics approval was sought from our local institutional review board and exemption was provided (2023/00104). Best efforts have been made to ensure anonymity.

Credit Author Statement

ST: Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review and Editing

AY: Data Curation, Validation, Visualization, Writing - Review and Editing

SSaw: Conceptualization, Resources, Supervision, Writing - Review & Editing

SSethi: Supervision, Writing – Review and Editing

Funding and Data availability statement

This case report received no specific grant from any funding agency in the public or commercial sectors. The data that support the findings are stored on a password-protected computer owned by our hospital due to sensitive patient information. Access to the data is restricted in accordance with institutional policies to protect patient privacy. Researchers interested in the data may contact the corresponding author, subject to institutional and ethical approvals."

References

1. Choy KW, Wijeratne N, Lu ZX, et al. Harmonisation of Osmolal Gap - Can We Use a Common Formula? Clin Biochem Rev. 2016;37:113–119.
2. Rasouli M, Kalantari KR. Comparison of methods for calculating serum osmolality: Multivariate linear regression analysis. Clin Chem Lab Med. 2005;43:635–640. doi: 10.1515/CCLM.2005.109
3. Smithline N. Gaps--anionic and osmolal. JAMA J Am Med Assoc. 1976;236:1594–1597. doi: 10.1001/jama.236.14.1594
4. Upadhyay A, Jaber BL, Madias NE. Incidence and Prevalence of Hyponatremia. Am J Med. 2006;119. doi: 10.1016/j.amjmed.2006.05.005
5. Cumming K, Hoyle GE, Hutchison JD, et al. Prevalence, incidence and etiology of hyponatremia in elderly patients with fragility fractures. PLoS One. 2014;9. doi: 10.1371/JOURNAL.PONE.0088272
6. Biswas M, Davies JS. Hyponatraemia in clinical practice. Postgrad. Med. J. 2007;83:373–378.
7. Rifai N, Horvath A, Wittwer C. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics 6th Ed. St. Louis, Mo: Elsevier 2018.
8. Chopra P, Datta SK. Discrepancies in Electrolyte Measurements by Direct and Indirect Ion Selective Electrodes due to Interferences by Proteins and Lipids. J Lab Physicians. 2020;12:84–91. doi: 10.1055/s-0040-1713690
9. Liu P. Blue book of diagnostic tests. Philadelphia: WB Saunders; 1986:239–254.
10. Unnikrishnan A, Pillai B, Pavithran P. Syndrome of inappropriate antidiuretic hormone secretion: Revisiting a classical endocrine disorder. Indian J Endocrinol Metab. 2011;15:208. doi: 10.4103/2230-8210.84870
11. Pathophysiology and etiology of the syndrome of inappropriate antidiuretic hormone secretion (SIADH) - UpToDate. https://www.uptodate.com/contents/pathophysiology-and-etiology-of-the-syndrome-of-inappropriate-antidiuretic-hormone-secretion-siadh?search=siadh&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1#H2 (accessed: 18/02/2025)
12. Hannon MJ, Thompson CJ. The syndrome of inappropriate antidiuretic hormone: Prevalence, causes and consequences. Eur. J. Endocrinol. 2010;162:5–12.
13. Mavani GP, DeVita M V., Michelis MF. A review of the nonpressor and nonantidiuretic actions of the hormone vasopressin. Front Med. 2015;2:19. doi: 10.3389/FMED.2015.00019
14. Luttinger D, Hernandez DE, Nemeroff CB, et al. Peptides and nociception. Int Rev Neurobiol. 1984;25:185–241. doi: 10.1016/S0074-7742(08)60680-7
15. Spasovski G, Vanholder R, Allolio B, et al. Clinical practice guideline on diagnosis and treatment of hyponatraemia. Eur. J. Endocrinol. 2014;170:1–47.
16. Badrick T. Evidence-based laboratory medicine. Clin Biochem Rev. 2013;34:43–46.
17. Bezuidenhout K, Rensburg MA, Hudson CL, et al. The influence of storage time and temperature on the measurement of serum, plasma and urine osmolality. doi: 10.1177/0004563215602028
18. Plebani M. The detection and prevention of errors in laboratory medicine. Ann Clin Biochem. 2010;47:101–110. doi: 10.1258/acb.2009.009222
19. Badrick T, Gay S, Mackay M, et al. The key incident monitoring and management system - history and role in quality improvement. Clin Chem Lab Med. 2018;56:264–272. doi: 10.1515/CCLM-2017-0219
20. Kattar M, Mei J, Cembrowski G. Assessment of Short-Term Individual Variation of Osmolality Based on Serial Patient Measurements. Am J Clin Pathol. 2016;146. doi: 10.1093/AJCP/AQW163.009
21. Jones GRD, Sikaris K, Gill J. Allowable limits of performance' for external quality assurance programs - An approach to application of the stockholm criteria by the RCPA quality assurance programs. Clin Biochem Rev. 2012;33:133–139.
22. Deetz CO, Nolan DK, Scott MG. An examination of the usefulness of repeat testing practices in a large hospital clinical chemistry laboratory. Am J Clin Pathol. 2012;137:20–25. doi: 10.1309/AJCPWPBF62YGEFOR
23. Osmometers | Advanced Instruments. https://www.aicompanies.com/osmometers/ (accessed: 20/02/2025)

Case Report

Unexplained Hypercobalaminemia: A Rare Diagnostic Pitfall

Anil K Chokkalla[1,2*], Kafi B Thomas[3], Arun Rathinam[4]

[1]*Laboratory Medicine, Sanford Health, Fargo, ND, USA*
[2]*Pathology, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND, USA*
[3]*Hematology and Medical Oncology, Sanford Cancer Center, Bismarck, ND, USA*
[4]*Hematology and Oncology, University of North Dakota School of Medicine and Health Sciences, Grand Forks, ND, USA*

Article Info

[*]***Corresponding Author:***
Anil Chokkalla, PhD, D(ABCC), NRCC
Clinical Chemist, Sanford Laboratories
Medical Director, Sanford Fargo Point-of-Care Testing
Clinical Assistant Professor, University of North Dakota
Office: 4C156, Sanford Broadway Medical Center
801 Broadway North, Fargo, ND-58102
E-mail: anilkiran.chokkalla@sanfordhealth.org

Keywords

Interference, Macro B12, Polyethylene Glycol, Immunoassay

Abstract

Persistently elevated serum vitamin B12 may suggest underlying pathology once analytical interferences are excluded. Here, we report a case of 56-year-old female with high vitamin B12 levels incidentally noted by primary care provider. Repeat testing with different assay method over time demonstrated consistently high values, prompting further evaluation. Hematology-oncology provider initiates malignancy workup due to well-documented association between neoplastic process and elevated vitamin B12 levels. Concurrently, clinical laboratory was contacted to rule out potential assay interference related to macro-vitamin B12 complexes. A polyethylene glycol precipitation study demonstrated markedly low vitamin B12 recovery, supporting the presence of a macro-vitamin B12 complex as the cause of spurious elevation. This case report highlights the importance of excluding macro–vitamin B12 interference prior to initiating extensive malignancy workup and the lack of evidence-based guidelines for the laboratory assessment of macro-complexes.

Case description

A 56-year-old female presents to the hematology-oncology clinic for the follow-up of elevated vitamin B12 levels. Past medical history is significant for hypothyroidism and being managed with levothyroxine (100 µg). At the time of initial encounter, patient was undergoing evaluation for fatigue, unintentional weight loss of 7 pounds over last couple months and headache without fever, vision changes and vomiting. Laboratory workup included complete blood count, peripheral smear examination, vitamin B12, and thyroid function panel. Patient was euthyroid with a free T4 of 1.4 ng/dL (reference interval: 0.7-1.5 ng/dL) and TSH of 2.71 µIU/mL (reference interval: 0.35-4.94 µIU/mL). Furthermore, complete blood count showed normal hemoglobin, white blood cells, absolute neutrophil and platelet counts. Patient's vitamin B12 levels are elevated at >1,500 pg/mL (reference interval: 180-914 pg/mL) in September 2025 at an outside facility using Siemens Atellica method. Repeat testing using same method yielded exact same result i.e. >1,500 pg/mL in October 2025. Primary care provider subjected this patient for vitamin B12 testing at our laboratory, which is based on Abbott Alinity method. Consistently, vitamin B12 levels are reported over upper limit of reportable range i.e. >2,000 pg/mL (reference interval: 213-816 pg/mL) on two different days that are one week apart. Patient denied taking supplements containing vitamin B12. Due to persistently elevated vitamin B12, hematologic malignancy was suspected by the hematology-oncology provider. As a part of malignancy workup, age-appropriate cancer screenings were reviewed. Patient's first colonoscopy at the age of 50 years showed a 2 mm sensile polyp in the cecum and 8 mm polyp in the ascending colon. Histopathology on resected tissue was notable for tubular adenoma without high-grade dysplasia. Patient's Esophagogastroduodenoscopy biopsy from 2020 indicated eosinophilic esophagitis. Patient's mammogram was benign with BI-RADS score of 1. Serum protein electrophoresis results were unremarkable with gamma fraction of 0.6 g/dL (reference interval: 0.5-1.3 g/dL). Physical exam did not reveal palpable lymphadenopathy or hepatosplenomegaly. Hepatic and renal function tests are all within normal limits. Evidence from literature shows association between elevated vitamin B12 levels and an underlying solid or hematologic malignancy such as hepatocellular, breast and colon carcinoma, chronic myeloid leukemia and polycythemia vera as well as renal and liver disease [1]. Patient's white blood count and absolute neutrophil count were normal, and differential with absence of basophilia and immature granulocytes. Review of peripheral blood smear by hematopathologist did not reveal any other abnormalities. CT scan (chest/abdomen/pelvis) was ordered for further evaluation. Collectively, the laboratory and clinical features do not support the presence of an underlying myeloproliferative neoplasm. Previous studies linked falsely elevated vitamin B12 levels due to assay interference secondary to macro-complex [2]. Provider contacted the clinical laboratory to assess the macro-vitamin B12 status

Case discussion

Vitamin B12 deficiency has been implicated in diverse pathological states such as neuropathy, cognitive dysfunction, anemia, cardiovascular disease, age-related macular degeneration, and gastrointestinal disorders. However, the clinical significance of elevated serum B12 levels is not well understood [3]. Elevated vitamin B12 levels often reflect recent supplementation, as serum concentrations typically peak about 6–8 hours after ingestion [4]. An estimated 10% laboratory specimens show vitamin B12 levels above the upper limit of normal [5]. Several preanalytical factors like 1) macro complexes, 2) heterophile antibodies and 3) autoantibodies against intrinsic factor are implicated in the literature [4]. Additionally, piecemeal evidence links elevated vitamin B12 levels to hematologic or solid-organ malignancies. High association was noted between hypercobalaminemia and neoplastic disease with a odds ratio of 1.8 [1]. Hence, Hematology/Oncology providers are presented with a dilemma to rule out preanalytical vs. clinical causes of elevated vitamin B12. The former is challenging due to lack of analytical methods, whereas the latter requires extensive workup. Circulating macro-complexes form predominantly due to binding of analyte or its carrier protein with immunoglobulins. Best known examples include macro-TSH, macro-amylase, macro-troponin, macro-prolactin and macro-creatine kinase. The key features of these macro-complexes are slow clearance from the bloodstream and lack of biological activity [5]. In the bloodstream, ~ 4/5th of vitamin B12 is bound to haptocorrin, while the 1/5th is carried by transcobalamin, the transport protein responsible for delivering vitamin B12 to tissues. Macro-vitamin B12 complexes are thought to originate from aberrant binding of carrier proteins with immunoglobulins or soluble receptors and in rare instances, from direct binding of free vitamin B12 with γ-globulins [4]. Polyethylene glycol (PEG) precipitation is a widely used non-specific method for the removal of macromolecules from plasma samples. Of note, gel filtration chromatography is the reference method but often limited to specialty reference laboratories. Currently, there are no evidence-based guidelines for troubleshooting macro-complexes.

To rule out interference due to macro B12 complex, we subjected the patient's specimen to PEG precipitation, followed by measurement of monomeric vitamin B12 in the supernatant as described previously [6]. Briefly, 100 µL of PEG 8000 solution (250 g/L) was mixed with 100 µL of patient plasma and incubated at 370C for 30 mins. The macro-vitamin B12 complex is precipitated by centrifuging at 13,000 rpm for 3 mins, and the supernatant is analyzed in duplicates along with neat sample. Vitamin B12 concentration of the pre-PEG treated sample was 1894 pg/mL, whereas the post-PEG treated sample was 169 pg/mL. Importantly, the recovery of vitamin B12 [2*(post-PEG B12/pre-PEG B12)x100] after PEG treatment was only 18% (Table 1). Previous studies recommend that a recovery of ≤50% after PEG precipitation is consistent with macro-vitamin B12 interference [3]. Hemotology-oncology

provider showed interest in elucidating the components of this macro-vitamin B12 complex. To further investigate the constituents of macro-vitamin B12 complex, the plasma levels of rheumatoid factor, immunoglobulin's G, A and M were measured. Interestingly, we did not notice any abnormality of these protein levels in the neat specimen (Table 1). It will be interesting to subject these specimens to confirmatory testing with gel filtration chromatography, but we did not have access to this method within our geographical location. It is plausible that soluble transcobalamin receptor is responsible for the macro-vitamin B12 complex formation [7]. Altogether, these findings are suggestive of potential interference with macro-vitamin B12 complex.

Table 1: PEG precipitation of macro-vitamin B12 complex.

Analyte	Result	Reference Interval
Vitamin B12 Neat (pg/mL)	>2000	213-816
	>2000	
Vitamin B12 Saline (pg/mL)	1853	
	1935	
Vitamin B12 PEG (pg/mL)	167	
	170	
Rheumatoid Factor (IU/mL)	<13	<30
Immunoglobulin G (mg/dL)	795	552-1631
Immunoglobulin M (mg/dL)	107	33-393
Immunoglobulin A (mg/dL)	314	65-421

Vitamin B12 is quantified on the Abbott Alinity platform using a two-step competitive chemiluminescent microparticle immunoassay. In this method, pre-treatment step releases vitamin B12 from carrier proteins which then competes with a labeled vitamin B12 analog for binding to intrinsic factor–coated paramagnetic microparticles, producing a chemiluminescent signal that decreases as the analyte concentration increases. Importantly, the assay package insert does not list macro-vitamin B12 interference as potential limitation. Clinical Laboratory Standards Institute guidelines do not currently provide specific recommendations for evaluating assay interference caused by macro-complexes. This gap leaves laboratories to rely on local protocols and published case reports when assessing potential macro-vitamin B12 interference. Even though vitamin B12 levels are consistently high due to macro-vitamin B12 complex, it is important to rule out functional deficiency using adjunct biomarkers such as methylmalonic acid, holo-transcobalamin and homocysteine, a notable limitation of the present case.

Ethical approval

Patient consent submitted to Sanford Health Information Management.

Declaration of conflicts

The authors declare no conflict of interest.

Funding and Data Availability

Not applicable.

CRediT author statement

Anil K Chokkalla: Conceptualization, Investigation, Writing – Original Draft; Kafi B Thomas: Conceptualization, Reviewing, Editing; Arun Rathinam: Reviewing, Editing.

References

1. Andrès E, Serraj K, Zhu J, Vermorken AJ. The pathophysiology of elevated vitamin B12 in clinical practice. Qjm. 2013;106(6):505-515.
2. Belkhouribchia J. Macro-Vitamin B12 as Cause of Falsely Elevated Cobalamin Levels. Eur J Case Rep Intern Med. 2024;11(1):004188.
3. Delgado JA, Pastor MI, Costa G, Márquez N, Bauça JM. Interference by vitamin B12-macrocomplexes: towards an effective detection and correct interpretation of hypo- and hypervitaminemia. Adv Lab Med. 2024;5(4):386-393.
4. Fedosov SN, Nexo E. Macro-B12 and Unexpectedly High Levels of Plasma B12: A Critical Review. Nutrients. 2024;16(5).

5. Duim SN, Vlasveld LT, Mezger STP, Mingels AMA, Ramakers CRB, de Boer D, et al. “Macro transcobalamin causing raised vitamin B12: Case-based laboratory investigation”. Ann Clin Biochem. 2022;59(4):302-307.
6. Anil K Chokkalla, David L Paul, Estella Tam, and Sridevi Devaraj. Immunoassay Interference due to Macro-TSH: A Case Study of a Pediatric Patient with Hashimoto’s Thyroiditis. Archives of Clinical and Medical Case Reports. 7 (2023): 124-128.
7. Hoffmann-Lücke E, Arendt JF, Nissen PH, Mikkelsen G, Aasly JO, Nexo E. Three family members with elevated plasma cobalamin, transcobalamin and soluble transcobalamin receptor (sCD320). Clin Chem Lab Med. 2013;51(3):677-682

Case Report

Proton Pump Inhibitor (PPI)-Domperidone Induced Anemia and Galactorrhea: A Case Report

Rupali Bains[1*], Sudhahar Tamizhan[1], Priyanka Patra[2], Uttam Kumar Nath[3], Manisha Naithani[4]

[1]*Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi, India*
[2]*Department of Clinical Genetics and Fetal Medicine, Kasturba Medical College (KMC), Mangalore, Karnataka, India*
[3]*Professor & Head, Department of Medical Oncology and Hematology, All India Institute of Medical Sciences (AIIMS), Rishikesh, Uttarakhand, India*
[4]*Professor, Deparmtent of Biochemistry, All India Institute of Medical Sciences (AIIMS), Rishikesh, Uttarakhand, India*

Article Info

[*]***Corresponding Author:***
Rupali Bains
Senior Resident, Department of Laboratory Medicine, All India Institute of Medical Sciences (AIIMS), New Delhi, India, 110029
E-mail: rupali.prinston@gmail.com

Keywords

Iron deficiency anemia-IDA, Proton pump inhibitors-PPIs, Ferric Carboxy Maltose-FCM, Fermented Oligo, Di, Monosaccharide and Polyols derivative-FODMAP Diet, Tissue transglutaminase-TTG

Abstract

Gastrointestinal disorders are increasingly common among young adults, particularly medical students exposed to chronic stress and irregular dietary habits. Proton pump inhibitors (PPIs) are widely prescribed for the management of gastroesophageal reflux disease (GERD), gastritis, and related acid-peptic disorders. However, prolonged PPI therapy has been associated with impaired iron absorption and may contribute to iron deficiency anemia (IDA).
We report the case of 26-year-old female, unmarried, postgraduate medical student with iron deficiency anaemia and galactorrhea. The patient presented to gastroenterology clinic with a complaint of mild non-colicky epigastric pain, nausea, fatigue and shortness of breath. Comprehensive evaluation included complete blood count with peripheral smear, iron profile study, stool routine and stool for occult blood, serum IgA anti TTG, viral markers, upper gastro intestinal endoscopy (UGIE) & biopsy, H. Pylori ammonia breath test, gastric emptying scintigraphy, USG abdomen and pelvis, and serum prolactin levels. Based on the clinical, laboratory, and radiological findings, a diagnosis of iron deficiency anemia secondary to long-term PPI use and domperidone-induced galactorrhea was established. The patient was treated with intravenous ferric carboxymaltose (1500 mg), advised discontinuation of PPI-domperidone therapy, and initiated on a low-FODMAP diet. Significant clinical and hematological improvement was observed on follow-up, along with resolution of galactorrhea. Although PPIs remain highly effective and generally safe agents for acid suppression, concerns regarding their long-term adverse effects are increasing. Iron deficiency anemia resulting from chronic PPI use and concurrent domperidone-induced galactorrhea occurring in the same patient is rarely reported. This case underscores the importance of meticulous history taking, appropriate diagnostic evaluation, and rational test utilization in identifying uncommon medication-related adverse effects. Clinicians should consider prolonged PPI use as a potential contributing factor when evaluating unexplained iron deficiency anemia, particularly in young individuals receiving long-term therapy.

Background

Iron deficiency anemia (IDA) is a major health issue in India for many reasons[1]. Long-term use of proton pump inhibitors (PPIs) has been associated with impaired iron absorption[2], while domperidone is known to cause hyperprolactinemia due to dopaminergic blockade[3]. The coexistence of these adverse effects is rarely reported. This case highlights on taking clinical history, ordering relevant diagnostic tests to rule out differential causes of IDA in young female and effective test utilization to avoid exhaustive investigations probably invasive GI biopsies.

Case presentation

A 26-year-old unmarried female young medical student presented to the Gastroenterology clinic with a complaint of mild non-colicky abdominal pain and nausea for 8 years. She had localized, dull, non-radiating epigastric pain started 6 years ago, present throughout the day and relieved on taking PPIs (Pantoprazole-Domperidone). She had a complaint of mild to moderate dyspnea on exertion, fatigue for past 3 years and history of white milky discharge from the nipple for the past 1 year. She had intermittent reflux once a week for the previous 9 years. The patient was on long term PPI (pantoprazole/ esomeprazole 40mg daily) and domperidone (30mg daily) combination from 2014 to 2021. For IDA, she was prescribed oral iron supplements in 2018, which she stopped taking after one week owing to non-compliance. Before presentation, she had a history of migraine with aura, for which the patient was on Sodium valproate + Propranolol prophylaxis for around 1.5 years (2017-2018). Valproate side effects had caused her to gain weight in the past. Before to presentation, she reported a 3kg intentional weight loss in 2021. The patient had no comorbidities. She had no history of surgery and menstrual irregularity. She is an occasional smoker, alcoholic, non-vegetarian, non-diabetic and non-hypertensive. She has no known drug allergy. There was no relevant family history in this report. On physical examination, the vital signs were as follows: Body temperature, 36.5°C; blood pressure, 118/72 mmHg; heart rate, 91 beats per min; respiratory rate, 20 breaths per min. There was no noticeable pallor, and the liver was just palpable below the rib. Patient's Laboratory workup, before and after taking treatment are shown in Table 1 below.

Table 1: Showing blood parameters before and after treatment for PPI-DA induced IDA and galactorrhea.

Blood parameters	Before treatment (2014-2021)	3 months after treatment (2022)
1. Hemoglobin (11.5 – 14.5 gm/dl)	**7.8**	**11.8**
2. Erythrocyte count (4.7 – 6.0 mil/cu.mm)	4.1	4.3
3. Packed cell volume (36 – 48 %)	26.1	37
4. Mean cell volume (78- 100 fl)	**60.0**	**81**
5. Mean cell hemoglobin (27 – 31 pg)	18.1	29.9
6. Mean cell Hb# conc. (32 – 36 g/dl)	28.9	33.2
7. Red cell distribution width (RDW) (12 – 14.5%)	21.0	21.5
8. Total Leucocyte count (4000 - 11000/cumm)	11200	10294
9. Platelet count (1.50 – 4.50 × 10^3/ul)	3.91	3.95
10. Peripheral smear	**Microcytic, hypochromic RBCs, anisocytosis seen**	**Normocytic, normochromic RBCs seen**
11. Stool for occult blood	Negative	Negative
12. Stool routine (ova/cyst/larvae/trophozoite segments of any parasites seen)	No	No
13. IgA Anti-TTG# (<4U/ml)	1.29	1.31
14. Iron (70-180ug/dl)	**6.0**	**78**
15. Ferritin (22-32ng/ml)	**2.2**	**24**
16. Vitamin B12 (156-672 pmol/L)	300.0	312.0
17. Folic acid (>5.38ng/ml)	6.63	6.78
18. Anti-HCV Ab (<0.90)	Non-reactive (0.02)	Non-reactive
19. Anti-HIV (<0.90)	Non-reactive (0.08)	Non-reactive
20. HBsAg# (<0.90)	Non-reactive (0.15)	Non-reactive

Blood parameters	Before treatment (2014-2021)	3 months after treatment (2022)
21. HELIC# Ammonia Breath test (<3.0 mm)	**Positive (+) (11.9mm)**	**Negative (-) (2.3mm)**
22. Follicle stimulating hormone (3.85-8.78 mIU/ml Follicular phase)	4.76	4.12
23. Luteinising hormone (2.12-10.89 mIU/ml Follicular phase)	5.22	6.30
24. Estradiol E2 (22.40-115.0 pg/ml Follicular phase)	36.54	37.57
25. Prolactin (2.80-29.20 ng/ml Non pregnant)	**94.01**	**21.02**
26. Inhibin B (21-53 pg/ml)	95.14	44.58

#Hb-Hemoglobin; TTG-Tissue transglutaminase; HELIC-Helicobacter Pylori; HBsAg-Hepatitis B surface antigen

USG abdomen and pelvis showed both ovaries bulky (right; 20cc, left; 18 cc) with echogenic stroma and no other significant abnormality. Upper GI endoscopy revealed a normal study up to duodenum D2, and a biopsy of the body and antrum revealed normal histology. Gastric emptying scintigraphy with 1mCi of Tc99m-Sulphur colloid delivered orally and serial imaging's recorded for 4 hours in both anterior and posterior views, showed no delays with gastric emptying time of t ½ 116 minutes (normal range 70-100 minutes) and 99% emptying after 4 hours.

Diagnosis and treatment

The final diagnosis, based on the patient's medical history and test results, was Chronic GERD/H. Pylori positive/ PPI-associated Iron deficiency anaemia and Domperidone-associated galactorrhea (drug-associated hyperprolactinemia). The patient was advised to avoid PPIs and Domperidone. Because the patient did not respond to oral iron therapy, Ferric Carboxy Maltose 1500 mg (1000 mg initial infusion followed by 500 mg after 10days) as per iron therapy standard dosing protocol, an H. Pylori kit course, folic acid 5mg and methylcobalamine 1500 mg OD were given for three months, and a FODMAP Diet; 1400kcal (low calorie) diet was recommended. On stopping the PPI-Domperidone combination therapy along with injection Ferric Carboxymaltose, IDA and Galactorrhea subsided completely. Follow-up serum iron studies, serum IgA, H. Pylori Ammonia breath test, and Glucose, Fructose, Lactose Hydrogen breath test were indicated, along with food adjustment and physical exercise for 30 minutes each day.

Discussion

Despite being safe for long-term use, Proton pump inhibitors (PPIs) can cause iron deficiency anemia[4]. Omeprazole-Domperidone was the first approved PPI for over-the-counter (OTC) use[5]. PPI-associated Galactorrhea, possibly caused by Domperidone, is rarely reported [6]. This case reported 27-year-old female postgraduate medical student who was taking PPI-domperidone daily for 8 years and experienced mild non-colicky abdomen pain and nausea and suggesting chronic development of iron deficiency.

Dietary iron contains both heme and non-heme iron. Heme iron ($Fe2^{+}$) is more easily absorbed once pancreatic enzymes release it from globin. Since ferric ($Fe3^{+}$) iron is predominant, it absorbs slowly[7]. Iron absorption from the duodenum requires STEAP3 or cytochrome b to reduce $Fe3^{+}$ to $Fe2^{+}$. Over-the-counter PPIs, such as esomeprazole, lansoprazole, and omeprazole, may lead to increased use and dependence. Proton pump inhibitors inhibit gastric H^{+}/K^{+}-ATPase in parietal cells, leading to hypochlorhydria. Gastric acid is essential for solubilizing dietary non-heme iron and reducing ferric (Fe^{3+}) to ferrous (Fe^{2+}) form via duodenal cytochrome b (Dcytb). Reduced gastric acidity impairs this conversion, thereby decreasing iron bioavailability and absorption in the duodenum. Intestinal reduction and absorption of $Fe3^{+}$ decreases. Gastric pH elevates G cell hyperplasia and serum gastrin. Due to a lack of available reduced ferric iron, ferrous iron absorption decreases, resulting in iron insufficiency. Additionally, increased gastric pH may impair release of iron from food matrices and reduce ascorbic acid-mediated reduction[8]. Long term PPI-associated IDA was first reported in 2002. According to Sharma et al. (2004), two patients who used PPI for severe gastritis developed iron deficiency that did not respond to oral iron until the PPI was stopped[9]. In another report, Sarazynski et al. found that 98 patients on PPI for more than a year had significantly lower haematocrit, hemoglobin concentration, and MCV than matched controls. Over 20% experienced a 1 g/dL reduction in hemoglobin concentration. These electronic medical record chart reviews were retrospective, therefore iron deficiency treatment was not discussed[10]. A 25-year-old PPI user in Japan experienced iron deficiency, according to Imai et al. in 2018, which was reverted back on stopping PPI. Anemia was linked with PPI usage in 51% of Japanese cardiac patients, compared to 19% without PPI use ($P < 0.001$). There were no conclusions other than that long-term PPI users may develop IDA[11].

On the other hand, one of the first reports of domperidone-associated galactorrhea came from Great Britain in 1983 and another came from India in 1991[12]. Domperidone is a peripheral dopamine D2 receptor antagonist. Dopamine

physiologically inhibits prolactin secretion from the anterior pituitary. Blockade of D2 receptors removes this inhibitory control, leading to hyperprolactinemia and consequent galactorrhea[13,14]. In our case, with relevant medical history and laboratory results, we found that our patient was a rare case, with coexistence of PPI-DA associated IDA plus galactorrhea. The Naranjo Adverse Drug Reaction Probability Scale has been evaluated for temporal association in this patient and a score of 8 derived[15] attached in Supplemantry file S01.

She had undergone several investigations in Table 1, before PPI- associated IDA was established as her diagnosis, which may be an exhausting experience. The hematological profile demonstrated classical iron deficiency anaemia with microcytosis (MCV 60 fL), hypochromia (MCH 18.1 pg), elevated RDW (21%), and depleted iron stores (ferritin 2.2 ng/mL). Normal B12 and folate levels excluded megaloblastic anaemia. Post-treatment normalization confirms iron deficiency as the primary etiology. Gastrointestinal bleeding excluded (negative stool occult blood, normal endoscopy), Celiac disease excluded (negative anti-TTG IgA), Nutritional deficiency unlikely (mixed diet), Helicobacter pylori was treated, yet anaemia persisted until PPI withdrawal supporting PPI as a contributing factor. Thyroid function was within normal limits. Further, prolactin was elevated (94 ng/mL) but normalized after domperidone withdrawal, and there were no symptoms (like headache, visual defects- absent) excluding pituitary pathology as cause of hyperprolactinemia. Rapid normalization after domperidone withdrawal strongly supports drug- associated etiology.

Finally, as part of her treatment plan, she was prescribed an injection of Ferric Carboxy Maltose 1500 mg (1000 mg initial infusion followed by 500 mg after 10days) as per iron therapy standard dosing protocol, an H. Pylori kit, a FODMAP diet, physical exercise and review after 3 months[16]. At 3-month follow-up Hb improved to 11.8 g/dl, Ferritin increased to 24 ng/mL, Prolactin normalized (21 ng/mL) and complete resolution of galactorrhea and fatigue was observed. A low FODMAP diet was advised to manage persistent functional gastrointestinal symptoms suggestive of IBS overlap, particularly bloating and dyspepsia, after exclusion of structural pathology[17]. The patient's summary of events and treatment history has been provided in Table 2.

Table 2: Showing Timeline of events and treatment course.

Year	Event	Treated
2014	GERD/Gastritis	Started Pantoprazole-Domperidone (*8 years daily)
2018	Anaemia	Oral Iron prescribed (1 week non-compliance stopped)
2020	Progressive fatigue	—
2021	Onset of Galactorrhea	—
2022	Diagnosis made Drug induced IDA with Galactorrhea	Ferric Carboxy Maltose 1500 mg (1000 mg initial infusion followed by 500 mg after 10days) H. Pylori kit course Folic acid 5mg and Methylcobalamine 1500 mg OD* 3 months, FODMAP Diet
2022	Causative drug PPI-Domperidone withdrawal	

This case highlights the importance of balancing comprehensive diagnostic evaluation with judicious test utilization. Although several investigations were normal, but they were helpful to exclude common causes of iron deficiency anemia and hyperprolactinemia, thereby strengthening the association with prolonged PPI-domperidone use. The case reinforces the value of detailed history-taking and systematic exclusion of alternative diagnosis when assessing uncommon drug-related adverse effects.

Conclusion

This case reported that a probable association between long-term PPI use and iron deficiency anaemia, as well as domperidone and hyperprolactinemia[18]. While causality cannot be definitively established from a single case, the temporal relationship and resolution after drug withdrawal support a probable link. Without therapeutic rationale, chronic and improper PPI use raises short- and long-term adverse effects. PPI- associated malabsorption is more common, physician should consider it when diagnosing iron deficiency

without other causes. Over the counter drugs becoming a daily diet are highly alarming serious health concern especially in adolescents, young adults and medical students.

Conflict of Interest

Authors declare no conflict of interest.

Patient Consent

Informed written consent was obtained from the patient for publication of clinical details.

CRediT statement

RB,ST- Conceptualization, Writing – original draft and Writing – review & editing; PP, UKN- Investigation, Methodology, Supervision, Validation, and Visualization.

Funding and Data Availability Statement

No funding is required for this study. No supplemental data available for this study.

Ethical Statement

Informed consent obtained from the patient for clinical history. This case report has been written in compliance with the ethical principles for medical research in accordance with the Declaration of Helsinki.

References

1. Felix F, Karun KM, Nagendraswamy C, Sadanandan DM, Damodaran Y, Barvaliya M, et al. Effect of Interventions on Iron-Deficiency Anemia Among School-Going Children in India: A Systematic Review and Network Meta-Analysis. J Res Health Sci. 2025 Jun 10;25(3):e00651. doi:10.34172/jrhs.8985 PubMed PMID: 40823948; PubMed Central PMCID: PMC12445883.(Accessed on 27 April 2026)
2. Wu YT, Lu YT, Chu CY, Chao HJ, Kuo LN, Cheng KJ, et al. Is use of a long-term proton pump inhibitor or histamine-2 receptor antagonist a risk factor for iron-deficiency anaemia in Taiwan? A neglected clinical drug-drug interaction. Fam Pract. 2025;42(2):cmad090. doi:10.1093/fampra/cmad090(Accessed on 27 April 2026)
3. Yedla D, Sharmila V. Atypical presentation of domperidone-induced galactorrhea. Indian J Pharmacol. 2022;54(5):381–382. doi:10.4103/ijp.ijp_351_22 PubMed PMID: 36537410; PubMed Central PMCID: PMC9846906. (Accessed on 27 April 2026)
4. Imai R, Higuchi T, Morimoto M, Koyamada R, Okada S. Iron Deficiency Anemia Due to the Long-term Use of a Proton Pump Inhibitor. Intern Med. 2018;57(6):899–901. doi:10.2169/internalmedicine.9554-17 PubMed PMID: 29151538; PubMed Central PMCID: PMC5891535. (Accessed on 27 April 2026)
5. Prikis M, MacDougall J, Narasimhadevara N. Proton Pump Inhibitor-Induced Galactorrhea in a Kidney Transplant Recipient: A Friend or Foe? Case Rep Transplant. 2020;2020:8108730. doi:10.1155/2020/8108730 PubMed PMID: 32566351; PubMed Central PMCID: PMC7254079. (Accessed on 27 April 2026)
6. Poovathingal MA, Bhat R, Ramamoorthi null. Domperidone induced galactorrhea: an unusual presentation of a common drug. Indian J Pharmacol. 2013;45(3):307–308. doi:10.4103/0253-7613.111913 PubMed PMID: 23833383; PubMed Central PMCID: PMC3696311. (Accessed on 27 April 2026)
7. Piskin E, Cianciosi D, Gulec S, Tomas M, Capanoglu E. Iron Absorption: Factors, Limitations, and Improvement Methods. ACS Omega. 2022;7(24):20441–20456. doi:10.1021/acsomega.2c01833 PubMed PMID: 35755397; PubMed Central PMCID: PMC9219084. (Accessed on 27 April 2026)
8. Chen TP, Wu MJ, Chen CH, Tsai SF. Bone health, renal outcomes, and iron deficiency anaemia in proton pump inhibitor versus histamine-2 receptor antagonist users: a retrospective cohort study based on TriNetX global collaborative network data. BMJ Open Gastroenterol. 2025;12(1). doi:10.1136/bmjgast-2024-001723 PubMed PMID: 10.1136/bmjgast-2024-001723. (Accessed on 27 April 2026)
9. Sharma VR, Brannon MA, Carloss EA. Effect of omeprazole on oral iron replacement in patients with iron deficiency anemia. South Med J. 2004;97(9):887–889. doi:10.1097/01.SMJ.0000110405.63179.69 PubMed PMID: 15455980. (Accessed on 27 April 2026)
10. Sarzynski E, Puttarajappa C, Xie Y, Grover M, Laird-Fick H. Association between proton pump inhibitor use and anemia: a retrospective cohort study. Dig Dis Sci. 2011;56(8):2349–2353. doi:10.1007/s10620-011-1589-y PubMed PMID: 21318590. (Accessed on 27 April 2026)
11. Boxer M. Iron deficiency anemia from iron malabsorption caused by proton pump inhibitors. EJHaem. 2020;1(2):548–551. doi:10.1002/jha2.96 PubMed PMID: 35845020; PubMed Central PMCID: PMC9175665. (Accessed on 27 April 2026)
12. Cann PA, Read NW, Holdsworth CD. Galactorrhoea as side effect of domperidone. Br Med J (Clin Res Ed). 1983;286(6375):1395–1396. doi:10.1136/bmj.286.6375.1395 PubMed PMID: 6404476; PubMed Central PMCID: PMC1547854. (Accessed on 27 April 2026)
13. Agrawal P, Gupta A, Goyal V, Raj A, Pandey S. Galactorrhoea: A Rare Side-effect of Domperidone. Journal, Indian Academy of Clinical Medicine. 2011;12(3):225–226. (Accessed on 27 April 2026)
14. Prasad DA, Agrawal DS, Pradhan DC, Rao DKM. DOMPERIDONE INDUCED GALACTORRHOEA IN REPRODUCTIVE AGE GROUP FEMALES- A CASE SERIES. J Popl Ther Clin Pharmacol. 2024;31(4):343–

346. doi:10.53555/jptcp.v31i4.5458(Accessed on 27 April 2026)
15. A method for estimating the probability of adverse drug reactions - Naranjo - 1981 - Clinical Pharmacology & Therapeutics - Wiley Online Library [Internet]. [cited 2026 Apr 29]. Available from: https://ascpt.onlinelibrary.wiley.com/doi/abs/10.1038/clpt.1981.154?sid=nlm%3Apubmed. (Accessed on 27 April 2026)
16. Nwagha UI, Nwagha TU. Intravenous ferric carboxymaltose for iron deficiency anaemia in pregnancy. The Lancet Global Health. 2024;12(10):e1567–1568. doi:10.1016/S2214-109X(24)00331-0 PubMed PMID: 39304228. (Accessed on 27 April 2026)
17. Ville A, McRae R, Nomchong J, Reidlinger DP, Davidson AR, Staudacher HM, et al. Effects of a Low FODMAP Diet in Inflammatory Bowel Disease and Patient Experiences: A Mixed Methods Systematic Literature Review and Meta Analysis. J Hum Nutr Diet. 2025;38(4):e70106. doi:10.1111/jhn.70106 PubMed PMID: 40765115; PubMed Central PMCID: PMC12326053. (Accessed on 27 April 2026)
18. Zamani NF, Sjahid AS, Tuan Kamauzaman TH, Lee YY, Islam MA. Efficacy and Safety of Domperidone in Combination with Proton Pump Inhibitors in Gastroesophageal Reflux Disease: A Systematic Review and Meta-Analysis of Randomised Controlled Trials. Journal of Clinical Medicine. 2022;11(18):5268. doi:10.3390/jcm11185268(Accessed on 27 April 2026)

Supplementary files

Supplementary Figure 1: Naranjo Adverse Drug Reaction Probability Scale

EVIDENCIO
MEDICAL DECISION SUPPORT

Algorithm-id: 661 | Version: 1.37 | Revision date: 2017-01-25

NARANJO ADVERSE DRUG REACTION PROBABILITY SCALE

The Naranjo ADR Probability Scale was developed to help standardize assessment of causality for all adverse drug reactions. The scale was also designed for use in controlled trials and registration studies of new medications, rather than in routine clinical practice.

Research authors: *Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C, Greenblatt DJ*

RESULT 2026-04-29 17:53

Score 8

Based on the following parameters:

1. Are there previous conclusive reports on this reaction?	Yes +1
2. Did adverse event appear after the suspected drug was given?	Yes +2
3. Did the adverse reaction improve when the drug was discontinued or a specific antagonist was given?	Yes +1
4. Did the adverse reaction appear when the drug was readministered?	Yes +2
5. Are there alternative causes that could have caused the reaction?	No +2
6. Did the reaction reappear when a placebo was given?	Not known or not done
7. Was the drug detected in any body fluid in toxic concentrations?	Not known or not done
8. Was the reaction more severe when the dose was increased, or less severe when the dose was decreased?	Not known or not done
9. Did the patient have a similar reaction to the same or similar drugs in any previous exposure?	Yes +1
10. Was the adverse event confirmed by any objective evidence?	Not known or not done

Disclaimer: This algorithm is provided for educational, training and information purposes. It must not be used to support medical decision making, or to provide medical or diagnostic services.

Publisher: IFCC Communications and Publications Division (IFCC-CPD)

The eJIFCC is a member of the **Committee on Publication Ethics (COPE).**

The eJIFCC (Journal of the International Federation of Clinical Chemistry) is an electronic journal with frequent updates on its home page. Our articles, debates, reviews and editorials are addressed to clinical laboratorians. Besides offering original scientific thought in our featured columns, we provide pointers to quality resources on the World Wide Web.

Published by:

ifcc
International Federation of Clinical Chemistry and Laboratory Medicine

www.ifcc.org