International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

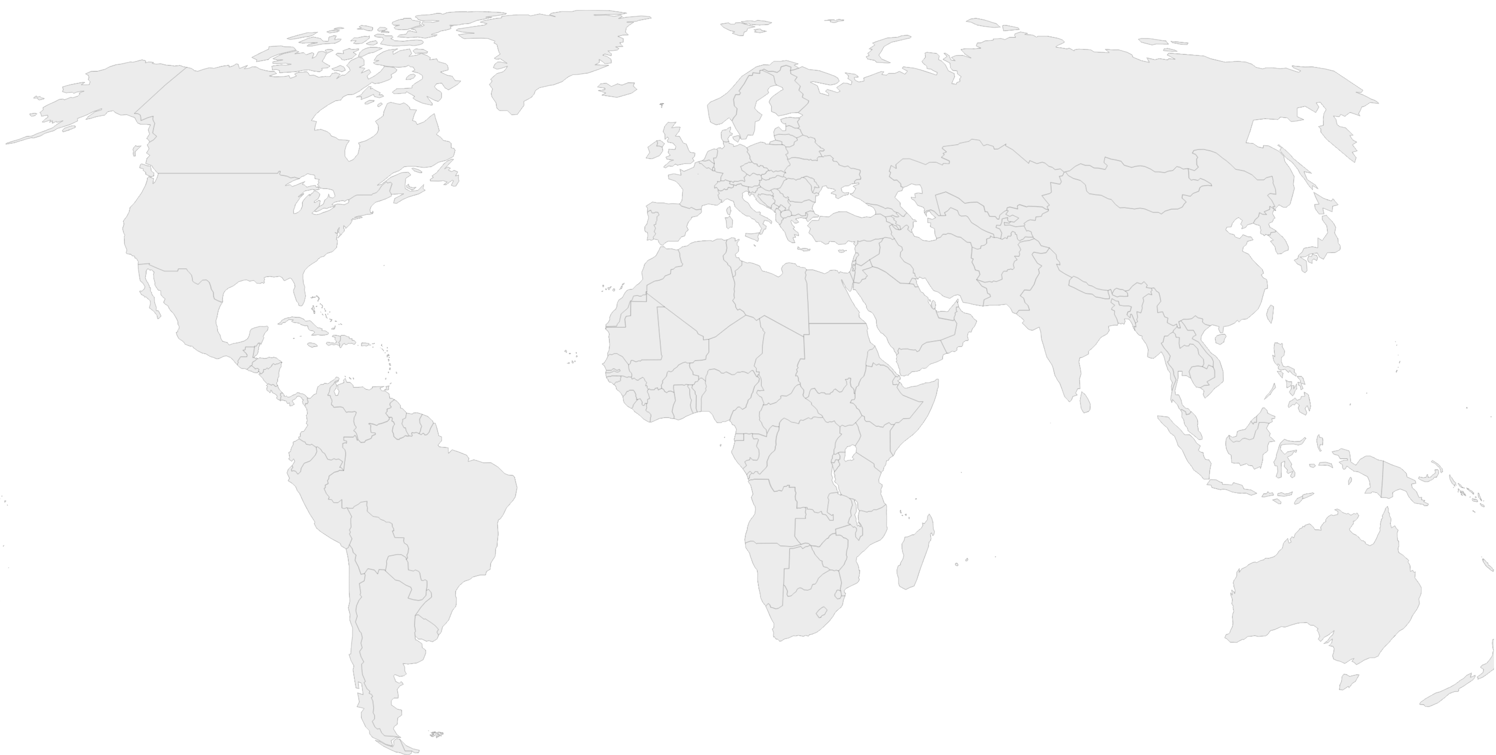
Communications and Publications Division (CPD) of the IFCC

June 2026
Volume 37
Number 3
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

Toward Standardization of Laboratory Terminology

Young Bae Lee Hansen, Fatma Meriç Yilmaz, Gunnar Nordin, Koh Furuta, Madeleen Bosma, Rebecca Ceder, Sridevi Devaraj — 389

Unraveling the need for standardization of Homocysteine assay: Insights from the results of External Quality Assurance

Sojit Tomo, Snigdha Singh, Amandeep Birdi, Manoj Khokhar, Dharmveer Yadav, Mithu Banerjee — 396

National Survey on Delta check Practices in Clinical Laboratories Across Pakistan

Fatima Kanani, Saba Raza, Harim Fatima, Adnan Mustafa Zubairi — 409

Big Data and Diabetes: Seven-Year Data from Pakistan's Nation-Wide Laboratory Network

Lena Jafri, Yousra Sarfaraz, Sibtain Ahmed, Zaib Un Nisa, Bilal Hashmi, Saad Bin Zafar Mahmood, Ayesha Sadiqa, Aysha Habib Khan, Imran Siddiqui, Hafsa Majid — 417

Bisalbuminemia: A rare finding identified via serum protein agarose gel electrophoresis and capillary electrophoresis techniques

Qian Ren, Tong Liu, Jing Li, Yulong Fan, Shoulei Wang, Lele Wang, Yansheng Wang, Zhaojing Liu, Yonghui Xia, Guoqing Zhu, Yansong Ren — 425

Polycystic Ovary Syndrome Prediction Using Machine Learning: A Comparative Analysis of Classification Algorithms

Usha Adiga, Vasishta Sampara, Pedda Reddemma P., Supriya P., Sireesha Kanchi, Kasala Farzia — 432

The Changing Landscape of Clinical Chemistry in Pakistan: Roles, Training Experiences, and Career Pathways- a National Survey

Sibtain Ahmed, Amal Mahmood, Aamir Ijaz, Nimarta Heman, Ayesha Sadiqa, Imran Siddiqui — 441

Assessing Testosterone Epidemiology Using Distributional Approaches: Evidence From Large-Scale Longitudinal Laboratory Dataset in Pakistan

Sibtain Ahmed, Minal Imran, Imran Siddiqui — 450

Comparative performance analysis of three point-of-care glucose testing devices against a central laboratory reference method

Auliya Ashar, Fauqa Arinil Aulia, Fauqa Arinil Aulia, Ferdy Royland Marpaung, Ferdy Royland Marpaung, M Robiul Fuadi, M Robiul Fuadi, Yessy Puspitasari, Yessy Puspitasari — 459

In this Issue

Autoantibody Profiles and Clinical Correlations in Systemic Sclerosis: A Cross-Sectional Study

Sushil KumarDebamita Datta, Debamita Datta, Sushil Kumar, Lekha Priyadharshini Kamarajan, Lekha Priyadharshini Kamarajan, Mala Mahto, Mala Mahto, Bandana Kumari, Bandana Kumari,, Ayan Banerjee, Swetalina Pradhan, Bijay Nanda Naik — 468

Advisory and Interpretative Services under ISO 15189:2022: A Prospective Study of Patient-Initiated Laboratory Consultations

Vivek Pant, Santosh Pradhan, Keyoor Gautam, Devish Pyakurel, Abha Shrestha, Neha Neupane — 476

Targeting Sample Reception Errors: A Failure Reporting Analysis and Corrective Action System (FRACAS)-Based Laboratory Quality Improvement Study

Vijetha Shenoy Belle, Merin K Thomas, Merin K Thomas, Nihaal Maripini, Nihaal Maripini, Saritha Kamath U ,Saritha Kamath U — 482

Integration of Patient-Based Real-Time Quality Control with Conventional Internal Quality Control: Improved Error Detection and Cost-Benefit Analysis in a Clinical Laboratory

Mudasir Bashir Dandroo ,Devanatha Desikan V,Ramesh Ramasamy — 493

The College of American Pathologists (CAP) Laboratory Accreditation Program (LAP): Accreditation Process and its Impact on Laboratory Practice and Patient Safety

Abdelaziz Sanad Abdelaziz, Samir Y. Marzouk, Samir Y. Marzouk, and Mai S. Mabrouk, and Mai S. Mabrouk — 506

A review of the methods available for the detection of antibodies against transglutaminase and deaminated gliadin in Celiac Disease, traditional and emerging technologies.

Diana Landoni, Gerson Dierley Keppeke — 517

Interrelationship between Serum FGF21, GDF15, and Microalbuminuria as Predictive Biomarkers for Early Detection of Diabetic Retinopathy in Type 2 Diabetes Mellitus

Mustakin, M.Kes, Sp.PK — 532

Acute Intermittent Porphyria with a Secondary Porphyria Cutanea Tarda-like Biochemical Pattern in a patient with co-morbid Human Immunodeficiency Virus infection

Kagiso M. Masemola, Kgaogelo R. Masemola Mogomotsi Dintshi, Mogomotsi Dintshi, Taryn Pillay, Taryn Pillay — 539

Monoclonal Gammopathy with Double M-bands Mimicking Biclonal Gammopathy: a case series

Dharmendra Kumar, Koushik Biswas, Kanu Tiwari, Mala Mahto — 545

In this Issue

The Case of Negative Low Density Lipoprotein Cholesterol 553

Earnest JP Daniel, Sridevi Devaraj

Drug-Induced Liver Injury During First-Line Anti-Tubercular Therapy (Isoniazid, Rifampicin, Pyrazinamide, and Ethambutol): A Case Report with De-challenge and Rechallenge 558

Sridevi Chokkakula Ashwitha Balasani, Sathwika Reddy Ammana

The Hypoxia Trap: Worsening Sickle Cell Disease in the Setting of Tetralogy of Fallot 566

Mohnish Darshan, Sonal Rajmane, Jignesh Sharma, Amber Kumar, Girish Chandra Bhatt, Shikha Malik

Anti-N-methyl-D-aspartate receptor autoimmune encephalitis due to ovarian teratoma in a 17-year-old young woman: case report and review of the literature 572

Ennio Polilli, Anna D'Andreagiovanni, Pierluigi Tocco, Maurizio Rosati, Ylenia Della Rocca, Giancarlo Di Iorio, Maria Vittoria De Angelis, Giustino Parruti, Antonella Frattari

Letter to the Editor

Toward Standardization of Laboratory Terminology

Young Bae Lee Hansen[1*], Fatma Meriç Yilmaz[2], Gunnar Nordin[3], Koh Furuta[4], Madeleen Bosma[5], Rebecca Ceder[6], Sridevi Devaraj[7]

[1]*Department of Clinical Biochemistry, North Zealand Hospital, Hillerød, Denmark*
[2]*Department of Clinical Biochemistry, Ankara Yildirim Beyazit University Medical Faculty, Turkey*
[3]*Equalis, Uppsala, Sweden*
[4]*Chiba Medical Center, Chiba, Japan*
[5]*Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, Leiden, The Netherlands*
[6]*Inera AB, Stockholm, Sweden*
[7]*Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX, USA*

Article Info

Corresponding Author:
Young Bae Hansen, PhD, Chair of C-NPU
Department of Clinical Biochemistry, North Zealand Hospital
E-mail: ysvl@hotmail.com

Keywords

Laboratory Information Systems, Standardization of terminology

Abstract

Background: Laboratory medicine is fundamental to evidence-based healthcare, providing critical data on human properties for diagnosing and managing diseases. The increasing complexity and size of laboratory data necessitate robust information technology (IT) systems for efficient management.

Purpose: This manuscript elucidates the roles and infrastructure requirements of laboratory IT systems by stating the purposes of laboratory medicine and examining the challenges and prerequisites associated with their achievement.

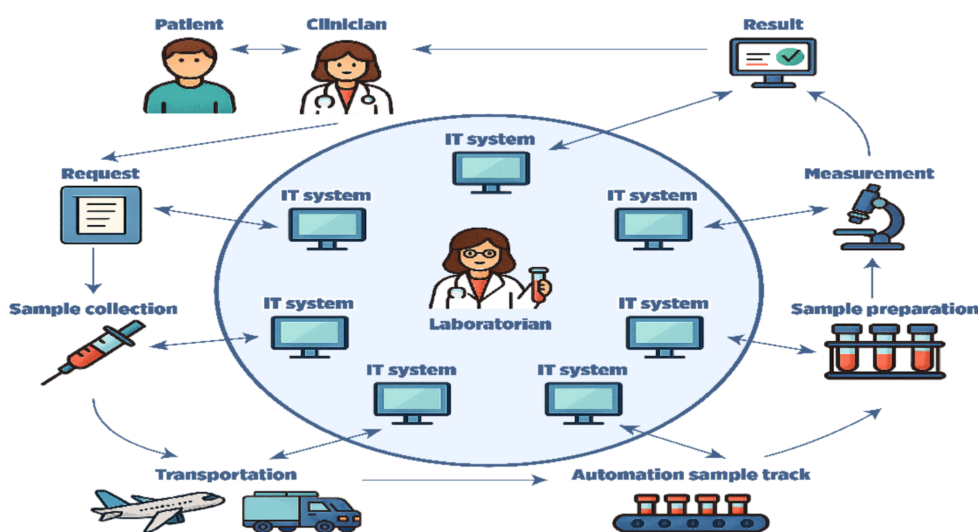
Methods: We explore the primary purpose of laboratory medicine in supporting clinical decisions and the opportunities for secondary data use. The challenges of data heterogeneity, insufficient and lack of metadata, and the need for standardized terminologies are addressed. We further emphasize the prerequisites including international collaboration and education.

Results: Technological advancements have enabled automation in laboratories, reducing errors, improving quality, and increasing efficiency and capacity for 24/7 sample processing. IT systems facilitate rapid results, enhance resource allocation, and manage the entire testing process from ordering to result interpretation. The establishment of two coherent and comprehensive laboratory information models - one for communication in the laboratory and another for communication between diverse health care providers - is proposed as essential for accurate data exchange and maximizing data utility.

Conclusion: Achieving successful exchange of laboratory results and optimal use of laboratory data hinge on international collaboration, shared resource commitment, and an approach to standardised terminologies and data structures. This work lays the foundation for ongoing efforts to develop international laboratory information models, critical for advancing patient care and leveraging the full potential of laboratory medicine.

Background
Medicine determines human properties to diagnose, monitor, predict, prevent, and treat diseases or injuries. Identifying an individual's biological and physiological properties is therefore fundamental to effective medical care. By recognizing patterns in these properties, clinicians can classify individuals into specific syndrome or disease categories, enabling targeted and personalized treatment strategies. Precise determination and interpretation of human properties form the cornerstone of evidence-based medicine.
The core function of laboratory medicine is to examine and measure an individual's chemical, biological, immunological, histological, and genetic properties through *in vitro* diagnostic (IVD) analysis of patient-derived materials. The vast number of laboratory tests performed globally each year reflects its central role in healthcare.
To ensure that the results of *in vitro* determinations accurately reflect the patients' true biochemical, physiological, or pathological conditions, not only the analytical steps but also the pre-and post-analytical steps (total testing process) must be upheld [1, 2]. Errors in any of these steps can compromise the result, obscure pathological processes, and potentially lead to improper patient treatment. With the developments of IT systems, laboratories may manage and control each aspect of the total testing process, thereby reducing these errors (Figure 1). It should be noted that structured information is required for machine-interpretability, decision algorithms and for proper training of AI systems.

Figure 1: Laboratory IT systems supporting the brain-to-brain loop.



The integration of total laboratory automation has dramatically increased laboratory efficiency and capacity, enabling rapid processing of a vast number of samples 24/7. This allows clinicians to receive test results shortly after sample collection, facilitating fast clinical decision-making. Furthermore, the lab information systems manage test requests, sample tracking, quality assessment, and transmit test results to Electronic Health Records (EHRs) for easy access by healthcare professionals and patients.
The integration of IT systems within laboratories has also created a foundational network facilitating digital connectivity between healthcare providers. First, this communication may provide real-time access to previous laboratory results from any health care provider involved in individual patient care at any given time and place. This is particularly critical in acute and life-threatening situations where immediate access to historical data can guide crucial treatment decisions, and it can also prevent unnecessary repeated examinations. Secondly, IT enables streamlined communication and eliminates the practice of commuting patients acting as carriers for their own laboratory information, often withcomplex technical terminology and nomenclature, between primary and secondary care and different healthcare organizations. Thirdly, the digital storage of laboratory results in extensive databases creates significant opportunities for data-driven (secondary) purposes.
This paper aims to elucidate the essential roles and infrastructure requirements of laboratory IT systems by stating the purposes of laboratory medicine and by examining the challenges and prerequisites associated with their achievement. Finally, we present a proposal for future directions on how to achieve the prerequisites.

Purposes of laboratory medicine
The primary purpose of laboratory medicine is to support informed clinical decisions for individual healthcare providers and patients. Laboratory requests and results should be appropriately presented to all relevant healthcare providers involved in individual patient care at any given place and time. Beyond this immediate clinical management, laboratory results, generated routinely for individual patient management, often accumulate in extensive databases. The substantial aggregated

data presents a significant opportunity for secondary use [3]. These secondary purposes include:

- Research and development [4].
- Providing evidence to support political healthcare decisions and clinical practice guidelines.
- Optimizing laboratory operations: enhancing laboratory production processes, for example through resource allocation and workflow management.
- Enabling local quality management: using patient data to assess analytical performance [5].
- Epidemiological surveillance: facilitating monitoring of drug utilization, antibiotic resistance, infectious disease trends, etc. [6].
- Administrative procedures: streamlining processes like billing and forecasting future laboratory workload.
- Conducting clinical value assessment: supporting evaluation of medical technologies, implemented procedures and their impact on patient outcomes [7].
- Directly enhancing patient health: identifying high-risk patient groups and supporting clinical value assessments that lead to improved care pathways [8].

In addition to high quality input data, the effective realization of these secondary purposes relies heavily on implementing appropriate IT tools capable of managing and analysing large datasets. The advent of IT tools, including AI, offers unprecedented capabilities for rapidly processing and interpreting this vast data [9, 10].

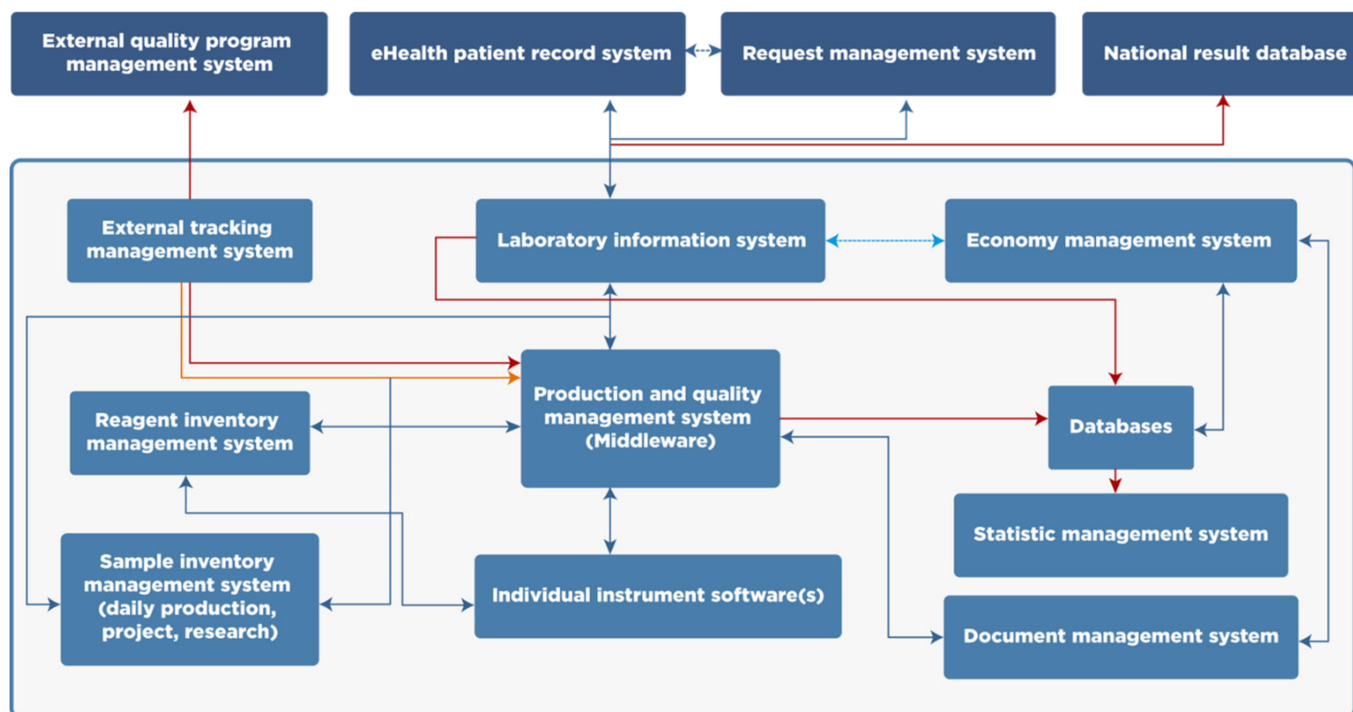
Challenges and Prerequisites for common laboratory data

Whether the goal is successful communication of individual patient results or accurate analysis of large laboratory datasets for secondary purposes, the underlying challenges and prerequisites are fundamentally the same. Accurate and consistent data exchange depends on the use of standardized, well-defined, and structured information that ensures that both the sender and the receiver interpret information in a shared and uniform manner. History has shown that this is not an easy task. Prior to the International System of Units (SI) results were reported with of hundreds of thousands of measurement units varying across countries, regions, towns and even districts [11, 12]. The number of units has decreased, but measurement units are still not standardized. Today, the existence of multiple "languages" for test requests and interpretation of test results arising from different cultures among healthcare providers and vendors of measuring systems, and adjusted to the needs of specialties,scientific fields and organization, continues to pose challenges. This heterogeneity can lead to communication errors, misinterpretation of results, and even patient harm, particularly when individuals transition between different hospitals or laboratory systems. For example, a clinician unfamiliar with the measurement unit used by another institution might misinterpret the unit for mass concentration of a critical quantity such as plasma digoxin instead of the unit for substance concentration [13]. Moreover, it requires cognitive actions to understand which one of the three units "µg/mL", "mg/L" and "mg/dL" that give numerical values that differ by a factor of 10, or to understand how to compare platelet counts if reported with the alternative units "cells per µL", "10^3/µL", "cells per nL", "10^9/L". Finally, Monjas et al. recently showed that despite the availability of UCUM, almost 60% of laboratory codes in routine healthcare data still use non-standard units, creating substantial obstacles for data interoperability and secondary use such as AI-based analytics or multicentre research [14].

Another crucial issue is the insufficiency or absence of essential information associated with laboratory test results. For example, lack of information about the method, calibration hierarchy, instrument type (and vendor) complicates the comparison of results from the same patient and property when different instrument types are used [15-17]. These inconsistent and insufficient information also hinder the effective secondary use of large datasets, necessitating time-consuming post-analysis data organisation and harmonization [18]. As calibration and procedures for the same property remain unharmonized or unstandardized, structured information is needed to address these differences to identify whether the results or data are comparable Braga et al. showed that, even after harmonisation efforts using successive generations of WHO reference materials for plasma ferritin, bias persisted among the various end-user measurement systems[19]. Until a robust solution is established, each result may be accompanied by information on the traceability of the quantity (and related metadata).

A practical consequence of these different data structures is the significant difficulty in exchanging data between different IT systems and platforms, a task that may seem technically feasible but is often complex and challenging. Data flow between existing, differently structured IT systems can be nearly impossible, and historical data may be lost during IT system transitions due to incompatible data structures. While a single IT system across all laboratories would be a solution, in practice, a uniform data exchange between interconnected IT systems from different laboratories may be a more feasible and scalable solution (Figure 2).

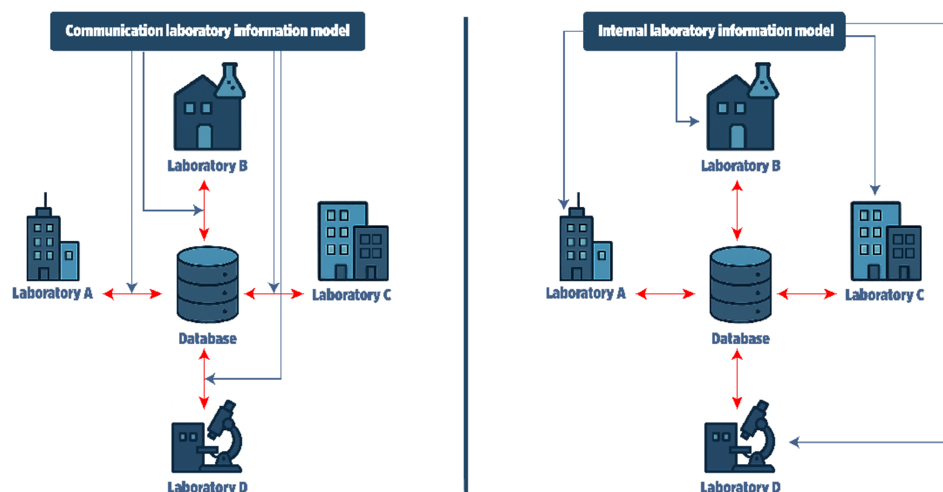
Figure 2: Data flow from different function in a future internal IT network.



The primary desired outcome is the establishment of an IT network and/or system that accurately reflects the entire laboratory process, from patient preparation to result interpretation. The second crucial outcome is the ability to present and communicate laboratory test results through IT networks to support correct clinical decision-making in any healthcare situation, at any place, and at any time. Furthermore, the collected data should provide support for laboratory automation, functions, artificial intelligence, and management.

To achieve these outcomes, two coherent, comprehensive, and information models should be established: one to support and communicate internal laboratory functions and another specifically for the communication and exchange of laboratory requests and results (Figure 3). The latter can be a subset of the first. Establishing these standardized models is essential for realizing the full potential of laboratory medicine in both clinical and analytical domains.

Figure 3: The communication laboratory information model is necessary for communication between different hospitals (left figure). Internal laboratory information model is the communication infrastructure within each laboratory (right figure).



Laboratory information model

A laboratory information model (also known as a laboratory data model/semantic data model/logical information model) serves as a dedicated "data structure" specific to laboratory medicine. It is characterized by "a representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse" [20, 21]. This model necessitates a structured content framework, incorporating internationally approved and standardized nomenclatures and terminologies [22], to ensure a meaningful and near-true representation of reality. The laboratory information model dictates the content and organization of laboratory databases. Consequently, a thorough description and analysis of desired outcomes' prerequisites are essential to identify relevant information types and propose content organization, structuring, and expression. Both the content and data model must possess the necessary granularity to provide sufficient and relevant information for achieving the particular outcome. Crucially, the structure and content of the models must be able to be revised to adapt to future advancements in science and technology.

Organization

Until a coherent laboratory information model has been agreed upon, the exchange of laboratory data across diverse organizational, national, and even temporal borders will remain a significant hurdle, potentially hindering collaborative research, public health initiatives, and the advancement of global healthcare.

The efforts to reach common understanding must be driven by international collaboration and coordination, actively engaging the full spectrum of stakeholders. Key participants should include experts in the laboratory, medical and terminology fields, while also representing international organizations (e.g., WHO, IFCC, EFLM, IUPAC, JCTLM (under BIPM), SNOMED CT, LOINC, ISO, etc.). IFCC could take a leading role in the coordination of this work. Furthermore, active involvement from laboratory IT, electronic health record providers and IVD companies is crucial to leverage technological expertise and ensure the practical implementation of the proposed laboratory information model. National governmental agencies and regulatory bodies must also be integral partners to facilitate the adoption and enforcement of agreed-upon frameworks within their respective jurisdictions.

The path from concept to international applicability includes four escalating phases:

Phase 1 – Definition

Agreement on core laboratory terminology and information model structure under the leadership of international scientific organizations. E.g., IFCC.

Phase 2 – Mapping

National terminologies are mapped to the master standard. Outputs include crosswalk tables usable by EHRs, LISs and IVD instruments.

Phase 3 – Pilot Deployment

Regional pilots (e.g., 3–5 countries or 20–50 laboratories) test live interoperability across systems.

Phase 4 – Progressive Adoption

Standardization

A prerequisite for achieving standardization (or at least harmonization) of laboratory result communication and data-driven initiatives is robust international collaboration. Engaging stakeholders, particularly the international laboratory communities, is paramount in establishing an accepted laboratory information model that ensures both quality and information relevance. Such international collaboration promotes a crucial sense of community and shared ownership, indispensable for reaching the international agreement (e.g., European Health Data Space (EHDS), FHIR (HL7) and openEHR [23-25]). This approach stimulates a common understanding of the challenges and proposed solutions, thereby securing widespread consensus and commitment.

A significant advantage of this collaborative framework is the avoidance of proprietary monopolies of laboratory information models, hindering interoperability. Establishing and continuously administering a shared organization and a comprehensive laboratory information model, while undeniably requiring resources and dedicated effort from all stakeholders, will ultimately impose a far smaller cumulative burden. A shared model inherently reduces the need for individual organizations to develop and maintain their own unique, potentially overlapping, data structures and communication standards, thereby minimizing redundant efforts in design, implementation, updates, and troubleshooting. Furthermore, it simplifies the training of laboratory personnel and IT professionals, as they only need to learn and manage a single model and a certain set of communication standards.

A central component of this shared model is the standardization of terminologies and nomenclatures. A lack of consistency leads to miscommunication - for instance, the ambiguous use of homonyms (same term, different meanings) and synonyms (different terms, same meaning) across publications and practice. These inconsistencies necessitate concerted international standardization efforts.

Achieving comprehensively standardized terminologies and nomenclature worldwide is a complex undertaking. Controversial concepts, e.g., gender/sex and ethnicity (biological inheritance), can pose significant challenges due to differing societal perspectives on the terms and their definitions. However, consistent, and carefully defined categories are essential for accurate clinical decision-making across patient categories. For instance, reference intervals for laboratory measurements may be specific to biological sex or ethnicity, directly impacting the accuracy and appropriateness of patient care. Therefore, reaching a consensus on the definitions of these fundamental concepts is paramount for ensuring consistent and appropriate healthcare delivery. Moreover, a well-designed and collaboratively maintained model, built upon standardized terminologies, can be more easily adapted to incorporate new technologies. Isolated systems, lacking this foundation, may struggle to keep pace with such advancements.

Conclusion

The path towards successful and rational exchange of laboratory results and optimal use of laboratory data hinge on robust international collaboration and coordination, a commitment to shared resources, and a meticulous approach to standardize terminologies and data structures. This initial letter lays the foundation for an ongoing series that will explore in greater detail the motivation, obstacles, and active efforts involved in building a functional proof of concept for international laboratory information models.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Gemini (Google) and LUMO (PROTON) in order to clarify and improve the readability and language of the manuscript. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Declaration of Conflict of interests

All authors are affiliated to the Committee on Nomenclature for Properties and Units (C-NPU), IFCC.

Ethical Approval

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

Funding

No funding to disclose.

Authors' contribution

Young Bae Hansen: Conceptualization, Writing – Original Draft, Visualization

Fatma Meriç Yilmaz: Conceptualization, Writing, Review, Editing, Visualization.

Gunnar Nordin: Conceptualization, Writing, Review, Editing,

Koh Furuta: Conceptualization, Writing, Review, Editing,

Madeleen Bosma: Conceptualization, Writing, Review, Editing,

Rebecca Ceder: Conceptualization, Writing, Review, Editing,

Sridevi Devaraj: Conceptualization, Writing, Review, Editing,

All authors have read and approved the final manuscript.

References

1. Armbruster D. Metrological Traceability of Assays and Comparability of Patient Test Results. Clin Lab Med. 2017;37(1):119-135. 10.1016/j.cll.2016.09.010.
2. Plebani M. Towards a new paradigm in laboratory medicine: the five rights. Clin Chem Lab Med. 2016;54(12):1881-1891. 10.1515/cclm-2016-0848.
3. Shirts BH, Jackson BR, Baird GS, Baron JM, Clements B, Grisson R, et al. Clinical laboratory analytics: Challenges and promise for an emerging discipline. J Pathol Inform. 2015;6:9. 10.4103/2153-3539.151919.
4. Lindvig KP, Thorhauge KH, Hansen JK, Kjaergaard M, Hansen CD, Johansen S, et al. Development, validation, and prognostic evaluation of LiverPRO for the prediction of significant liver fibrosis in primary care: a prospective cohort study. Lancet Gastroenterol Hepatol. 2025;10(1):55-67. 10.1016/S2468-1253(24)00274-7.
5. Kazmierczak SC. Laboratory quality control: using patient data to assess analytical performance. Clin Chem Lab Med. 2003;41(5):617-27. 10.1515/CCLM.2003.093.
6. Jian SW, Chen CM, Lee CY, Liu DP. Real-Time Surveillance of Infectious Diseases: Taiwan's Experience. Health Secur. 2017;15(2):144-153. 10.1089/hs.2016.0107.
7. Sikaris KA. Enhancing the Clinical Value of Medical Laboratory Testing. Clin Biochem Rev. 2017;38(3):107-114.
8. VanNess R, Swanson KM, Grenache DG, Koenig M, Dozier L, Freeman A, et al. Leveraging longitudinal clinical laboratory results to improve prenatal care. Am J Manag Care. 2021;27(2):60-65. 10.37765/ajmc.2021.88582.
9. Gruson D, Helleputte T, Rousseau P, Gruson D. Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulation. Clin Biochem. 2019;69:1-7. 10.1016/j.clinbiochem.2019.04.013.
10. Padoan A, Plebani M. Flowing through laboratory clinical data: the role of artificial intelligence and big data. Clin Chem Lab Med. 2022;60(12):1875-1880. 10.1515/cclm-2022-0653.
11. Alder K. The Measure of All Things: The Seven-year Odyssey and Hidden Error that Transformed the World: Free Press; 2002. 9780743216753.
12. Gyllenbok J. Encyclopaedia of Historical Metrology, Weights, and Measures : Volume 1. Cham: Springer International Publishing : Imprint: Birkhäuser,; 2018.
13. Flatman R. Terminology, units and reporting - how harmonized do we need to be? Clin Chem Lab Med. 2018;57(1):1-11. 10.1515/cclm-2017-1083.
14. Munoz Monjas A, Rubio Ruiz D, Perez Del Rey D, Palchuk MB. Enhancing real world data interoperability in healthcare: A methodological approach to laboratory unit harmonization. Int J Med Inform. 2025;193:105665. 10.1016/j.ijmedinf.2024.105665.
15. Antonsen S, Amundsen EK, Ceder R, Toska K, Tollanes MC, Hansen YB, et al. History, Implementation and Current Use of the IFCC-IUPAC's Nomenclature for Properties and Units (NPU) Terminology in Denmark, Norway and Sweden. EJIFCC. 2024;35(3):154-165.
16. Padoan A, Cadamuro J, Frans G, Cabitza F, Tolios A, De Bruyne S, et al. Data flow in clinical laboratories: could metadata and peridata bridge the gap to new AI-based applications? Clin Chem Lab Med. 2025;63(4):684-691. 10.1515/cclm-2024-0971.
17. Bosma M, Cobbaert C. Digital metrology in laboratory medicine: a call for bringing order to chaos to facilitate precision diagnostics. Clin Chem Lab Med. 2025;63(8):1486-1493. 10.1515/cclm-2025-0152.
18. Kim M, Shin SY, Kang M, Yi BK, Chang DK. Developing a Standardization Algorithm for Categorical Laboratory Tests for Clinical Big Data Research: Retrospective Study. JMIR Med Inform. 2019;7(3):e14083. 10.2196/14083.
19. Braga F, Pasqualetti S, Frusciante E, Borrillo F, Chibireva M, Panteghini M. Harmonization Status of Serum Ferritin Measurements and Implications for Use as Marker of Iron-Related Disorders. Clin Chem. 2022;68(9):1202-1210. 10.1093/clinchem/hvac099.
20. Lee YTT. Information Modeling: From Design to Implementation. IEEE Transactions on Robotics and Automation 1999 https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=821265

21. Chen PP-S. The entity-relationship model—toward a unified view of data. ACM Trans Database Syst. 1976;1(1):9–36. 10.1145/320434.320440.
22. Strike PW, Michaeloudis A, Green AJ. Standardizing clinical laboratory data for the development of transferable computer-based diagnostic programs. Clin Chem. 1986;32(1 Pt 1):22-29.
23. Shabani M. Will the European Health Data Space change data sharing rules? Science. 2022;375(6587):1357-1359.1126/science.abn4874.
24. Stellmach C, Muzoora MR, Thun S. Digitalization of Health Data:Interoperability of the Proposed European Health Data Space. Stud Health Technol Inform. 2022;298:132-136. 10.3233/SHTI220922.
25. REGULATION (EU) 2025/327 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, (2025). https://eur-le x.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202500327

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Unraveling the need for standardization of Homocysteine assay: Insights from the results of External Quality Assurance

Sojit Tomo[1], Snigdha Singh[1], Amandeep Birdi[1], Manoj Khokhar[1], Dharmveer Yadav[1], Mithu Banerjee[1*]

[1]*Department of Biochemistry, All India Institute of Medical Sciences, Jodhpur, Rajasthan, India*

Article Info

Corresponding Author:

Mithu Banerjee MBBS, MD, MNAMS
Professor & Head Department of Biochemistry
All India Institute of Medical Sciences Jodhpur, Rajasthan, India
Phone: +919622277766
E-mail: mithu.banerjee.3@gmail.com

Keywords

Assay standardization, External Quality Assurance Services, Homocysteine, Inter-platform variability, Limit of detection, NIST SRM 1955, Standard Reference Material, Traceability of calibrators

Abstract

Homocysteine is an intermediate product of the biosynthesis of methionine and cysteine and a sulfhydryl-containing amino acid. The study aims to assess the difference between the estimated levels of homocysteine by different automated systems on Biorad External Quality Assurance Services samples and the imprecision of examination of homocysteine levels below the limit of detection in External Quality Assurance Services samples by proficiency testing providers. This study involves the analysis of Biorad External Quality Assurance Services sample reports of two cycles (cycles 10 and 11) for homocysteine level evaluation performance. Our study demonstrates the wide difference in the bias percentage of homocysteine levels when bias was calculated with the peer mean and the method mean. There was a substantial increase in bias percentage for the Siemens Advia Centaur XP machine for samples with levels of homocysteine below the limit of detection. Our study also demonstrated that Siemens Advia and Siemens Atellica were evidently underestimating homocysteine levels when compared with other platforms. The use of Standard Reference Material for traceability of calibrators will ensure the standardization of assays between different manufacturers. The variation observed in the External Quality Assurance Services results of homocysteine highlights the impending requirement of standardization of assays using Standard Reference Materials.

Introduction

Homocysteine is an endogenously derived intermediate product of methionine metabolism. The normal range of homocysteine is from 5 to 15 µmol/L [1]. Human plasma contains homocysteine in three forms: free or unbound, protein-bound, and homocysteine-cysteine or homocystine dimers. Homocysteine level in humans is regulated by three enzymatic pathways based on the metabolic status. The various genetic factors that determine homocysteine levels in an individual include homozygosity or heterozygosity for Cystathionine-beta Synthase defects or Methylene Tetrahydrofolate Reductase defects, cobalamin mutations, and Down's syndrome [2,3]. Apart from genetic factors, physiological determinants such as gender, aging, decreased renal function, increased muscle mass, and lifestyle determinants such as habit of smoking, coffee, or ethanol intake also affect the homocysteine levels [4-6]. Clinical conditions such as folate deficiency, vitamin B12 deficiency, vitamin B6 deficiency, hypothyroidism, hyperproliferative disorders, and drugs such as folate antagonists, antiepileptics, contraceptives, vitamin B12 and B6 antagonists, aminothiols, etc also play a role in determining homocysteine levels [7]. Among the various determinants, only a few are responsible for lowering the level of homocysteine, such as the use of contraceptive pills or hormone therapy, aminothiols (penicillamine, acetylcysteine), or clinical conditions such as Down's syndrome [8,9]. The majority of the determinants increase the level of homocysteine. Hyperhomocysteinemia can be defined as plasma homocysteine levels of more than 15 µmol/L. Mild hyperhomocysteinemia is defined as the level of homocysteine between 15 to 30 µmol/L, moderate as the level of homocysteine between 30 to 100 µmol/L, and severe as the level of homocysteine greater than 100 µmol/L [10]. Hyperhomocysteinemia can clinically indicate cardiovascular, neurological, and psychiatric pathology. The mechanism underlying hyperhomocysteinemia contributing to cardiovascular, neurological, or psychiatric pathology points towards oxidative stress, deoxyribonucleic acid damage, protein homocysteinylation, or protein thiolation [11-13]. Homocysteine is auto-oxidized in plasma, leading to the formation of homocysteine disulfides, homocysteine, and homocysteine thiolactones. The initiation of lipid peroxidation is triggered by superoxide anion radical, hydrogen peroxide, and hydroxyl radical generated during the oxidation of homocysteine, which is ultimately responsible for the endothelial cytotoxicity of homocysteine. It has also been shown that a higher concentration of homocysteine restricts the activity of the antioxidant enzyme glutathione peroxidase [10,14,15]. Studies have shown that hyperhomocysteinemia induces inflammation in the mouse retina, brain, cultured retinal and microglial cells, and hence, targeting the reduction of inflammation is essential for mitigating damage associated with hyperhomocysteinemia age-related disorders such as diabetic retinopathy, age-related macular degeneration, and Alzheimer's disease [16]. It has also been shown that prenatal hyperhomocysteinemia in mother rats can induce behavioural impairments and oxidative stress in offspring rats. Some studies correlate the level of homocysteine with disease susceptibility [14]. A meta-analysis by Boushley et al had shown that the odds ratio of coronary artery disease in men with a 5 µmol/L increase in total homocysteine levels is 1.6, and that in women it is 1.8, whereas the odds ratio of cerebrovascular disease with an increase of 5 µmol/L in total homocysteine levels is 1.5. Hence, a change of 5 µmol/L is a clinically significant change [17]. The evaluation of homocysteine level is also important in neonates in diagnosing hemolytic uremic syndrome subtype and infantile tremor syndrome [18,19]. In the pediatric population, the elevated serum homocysteine levels correlate with autistic spectrum disorder and attention deficit hyperactivity disorder [20]. Our lab estimates homocysteine using the chemiluminescence method on the Siemens Advia Centaur XP platform. The assay range for the kits used in Siemens Advia Centaur XP is from less than 0.50 to 65 µmol/L. The quality control of the homocysteine parameter is performed by three levels of control samples. The range of level 1 is 4.03 to 8.16 µmol/L, whereas the range of level 2 is from 7.18 to 13.4 µmol/L, and that of level 3 is from 16.50 to 28.80 µmol/L. The Advia homocysteine assay is traceable to an internal standard manufactured using highly purified materials. The lab is enrolled in the Biorad External Quality Assurance Services cardiac markers programme for proficiency testing. Interestingly, on analysis of Biorad External Quality Assurance Services reports, we observed that multiple samples for homocysteine have been reported below the limit of detection of the assay. To our knowledge , no evidence was found in the literature to explain such variation of Homocysteine. Therefore, we aimed to analyze the Biorad External Quality Assurance Services sample reports of homocysteine for two cycles (cycles 10 and 11) and compare the homocysteine levels between different manufacturers.

Materials and methods

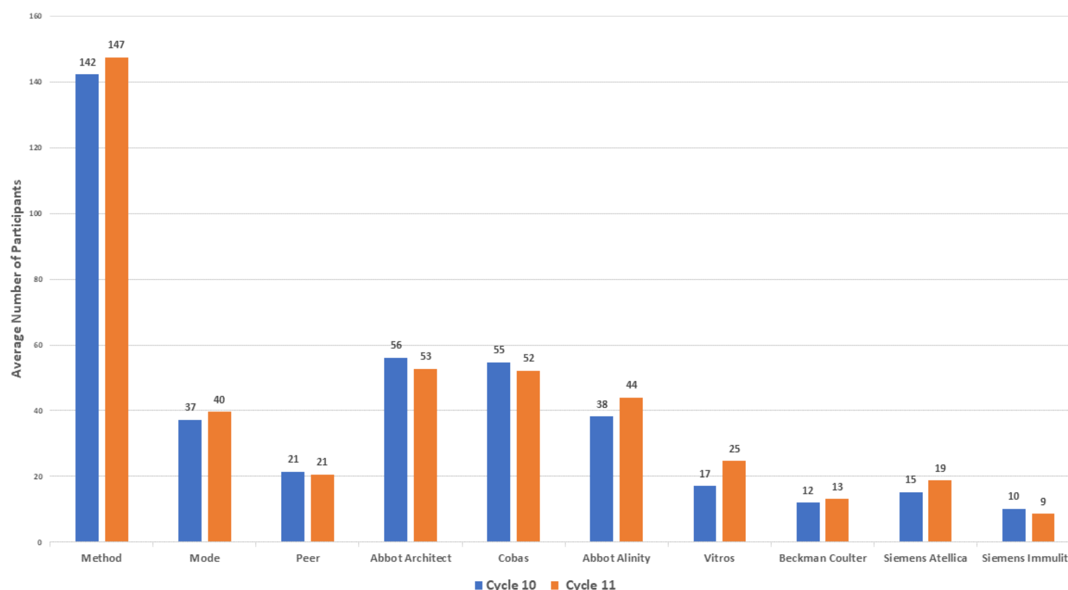
This is an observational study containing the External Quality Assurance Services reports of two cycles of the cardiac markers program of Biorad (cycles 10 and 11). The study includes the analysis of homocysteine levels of 24 External Quality Assurance Services samples evaluated using seven different machines: Abbot Architect, Cobas, Abbot Alinity, Vitros, Beckman Coulter, Siemens Immulite, Siemens Advia, and Siemens Atellica. Each cycle consists of 12 samples. Each new sample was reconstituted freshly and analyzed with the help of the analyzer. After this, the evaluated result was submitted to the Biorad External Quality Assurance Services panel for further analysis. In the graphical representations used for the demonstration of analysis of data, the cycle-10 samples 1 to 12 are numbered as 1 to 12, whereas the cycle-11 samples 1 to 12 are numbered as 13 to 24. The graph plots are designed based on the data in the report using MS Excel version 16.53 and SPSS version 23.

This study involved analysis of de-identified proficiency testing data without access to patient information or samples. No ethical approval was required for the analysis of anonymized quality control data according to institutional policies. All mandatory laboratory health and safety procedures were complied with during the course of conducting experimental work. Standard laboratory safety protocols were followed for handling and reconstitution of External Quality Assurance Services samples.

Results

The study includes data from around 142 labs in cycle 10 and 147 labs in cycle 11 as the method group, 37 labs in cycle 10 and 40 labs in cycle 11 as the mode group, and 21 labs each in cycle 10and cycle 11 as the peer group. The study incorporates the data on the number of labs using different platforms for the evaluation of homocysteine levels in the Biorad External Quality Assurance Services program, illustrated in Figure 1.

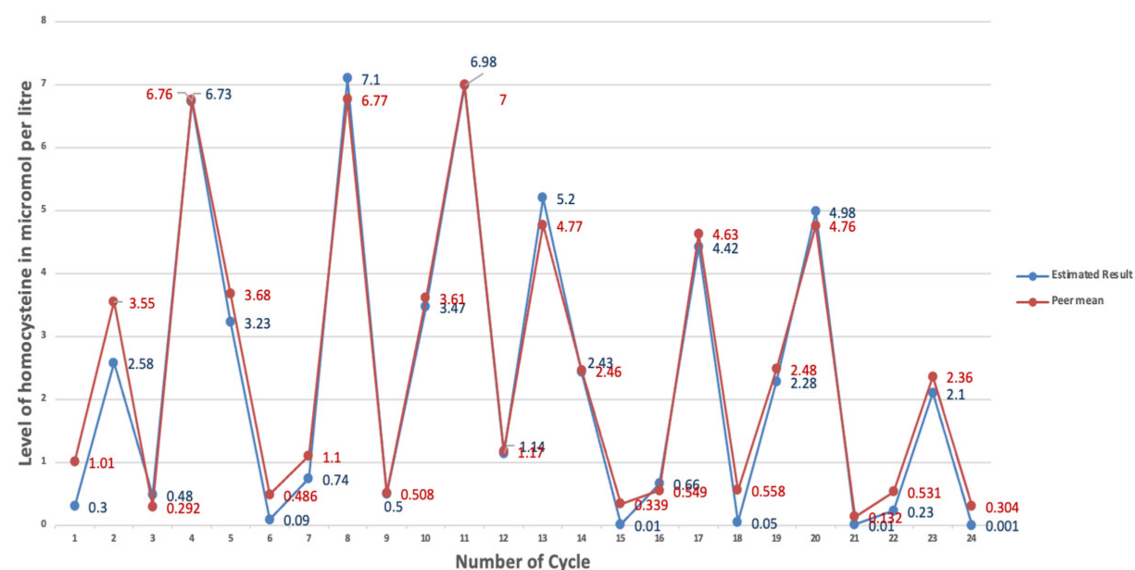
Figure 1: Distribution of Participating Laboratories Across Analytical Platforms.



Bar chart showing the number of laboratories utilizing each of the seven analytical platforms in the Bio-Rad EQA Program. Abbott Architect was the most commonly used platform (n=45), followed by Siemens Advia Centaur XP (n=21), Cobas (n=18), and Beckman Coulter (n=15). Data aggregated from Cycles 10 and 11.

The comparison of the mean serum homocysteine levels of 24 Biorad External Quality Assurance Services samples of the lab and the peer group is shown in Figure 2. The figure aptly shows that the lab mean was almost always 2 Z-scores when compared to the peer mean. The normal range of serum homocysteine in humans is from 5 µmol/L to 15 µmol/L. Among the 24 mean values of the peer group, only three values lie within the human serum range. The rest of the values are below the normal range. Among all the mean values, five values are even below the analytical measurement range of the kit.

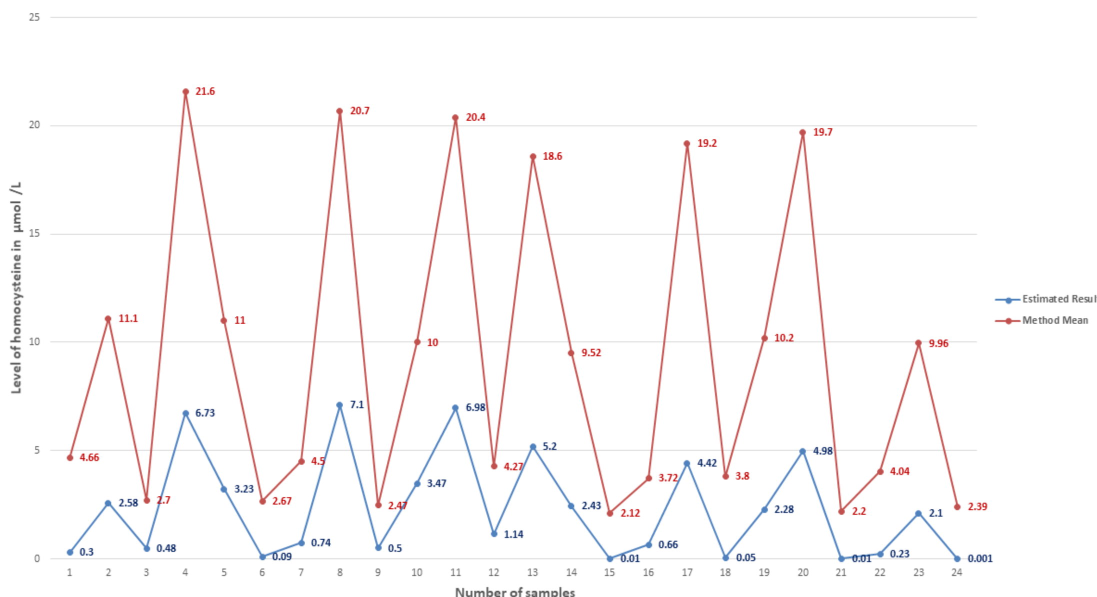
Figure 2: Comparison of Laboratory Results with Peer Group Means Across 24 EQA Samples.



Line graph comparing our laboratory's homocysteine measurements (Siemens Advia Centaur XP) with peer group means for all 24 samples (Cycle 10: samples 1–12; Cycle 11: samples 13–24). The laboratory consistently underestimated homocysteine, with most values yielding Z-scores <−2.0. Horizontal reference lines indicate the physiological range (5–15 µmol/L) and platform LOD (0.5 µmol/L). Only 3 of 24 peer group means fell within the physiological range.

The observation of the lab mean and the method mean reveals a wide difference between the lab mean and the method mean, as shown in Figure 3. Method mean refers to the mean value of all the labs that are using the same method or principle of evaluation for that analyte. It is interesting to note that the variation in the evaluation of homocysteine using the same method on different analyzers is wide.

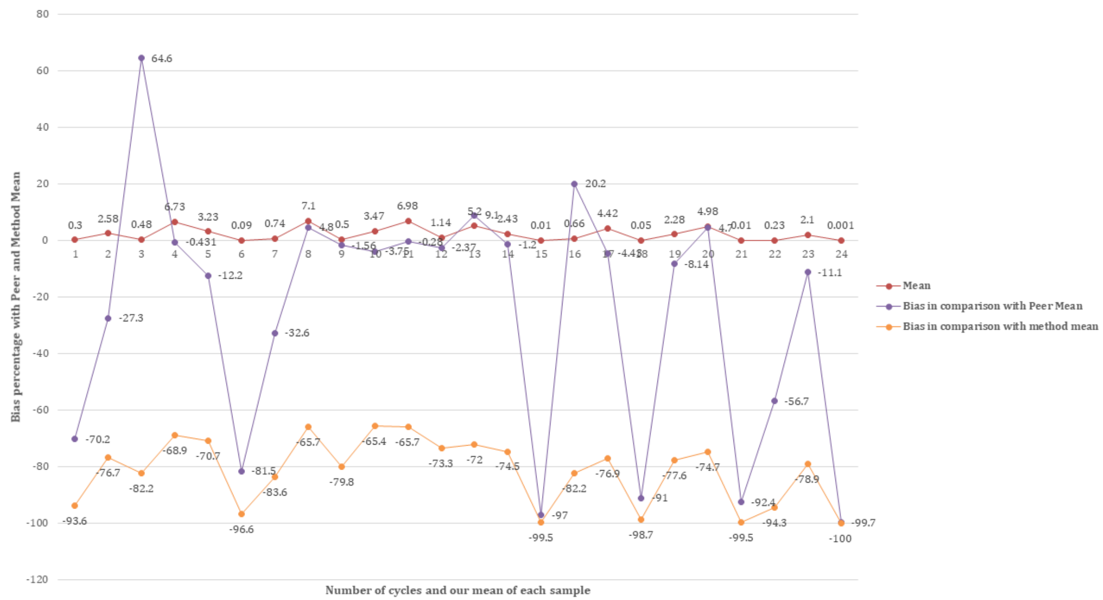
Figure 3: Laboratory Mean Versus Method Group Mean Discordance.



Line graph illustrating the difference between laboratory results and method group means (all chemiluminescent immunoassay platforms) across 24 EQA samples. Substantial discordance is evident, with differences exceeding 50% for low-concentration samples (<2 µmol/L), highlighting inter-analyzer variability even among platforms using ostensibly similar analytical principles.

The bias percentage of our lab mean is also compared with the peer mean and with the method mean. In Figure 4, it is clearly shown that the bias with the peer mean is high whenever the mean value is below the limit of detection of the Siemens homocysteine kit (0.5 µmol/L), due to the inherent imprecision. In the figure, sample numbers 1, 3, and 6 of cycle-10 and samplenumbers 3, 6, 9, and 12 of cycle-11 (noted as 15, 18, 21, and 24, respectively, in the graph) are either below the limit of detection or at the limit of detection. It is very obvious to note that there is a wide difference in the bias percentage when bias was calculated with the peer mean and the method mean.

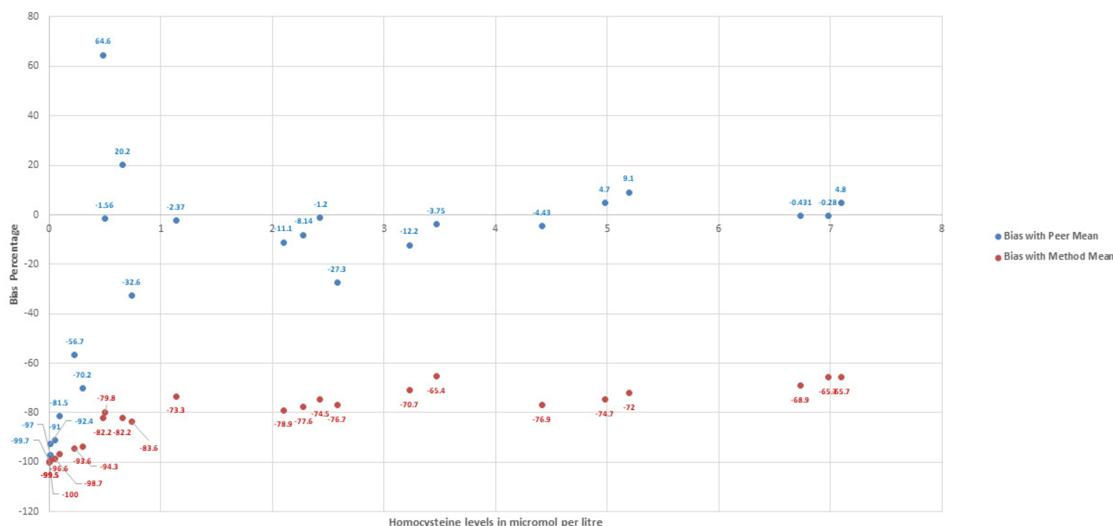
Figure 4: Bias Percentage: Peer Group Versus Method Group Comparison.



Dual-axis line graph comparing bias percentages calculated using peer group means (blue line) versus method group means (orange line) for all 24 samples. Samples below or near the LOD (0.5 µmol/L) - specifically samples 1, 3, 6, 15, 18, 21, and 24 - exhibit dramatically inflated bias percentages (−50% to −80%), demonstrating the impact of measurements at the lower analytical limit.

The scatter plot in Figure 5 shows the increased bias percentage of the Siemens Advia Centaur XP machine at low levels of homocysteine. The plot also reveals a continuously high biaspercentage on comparison of our lab values with the peer and the method.

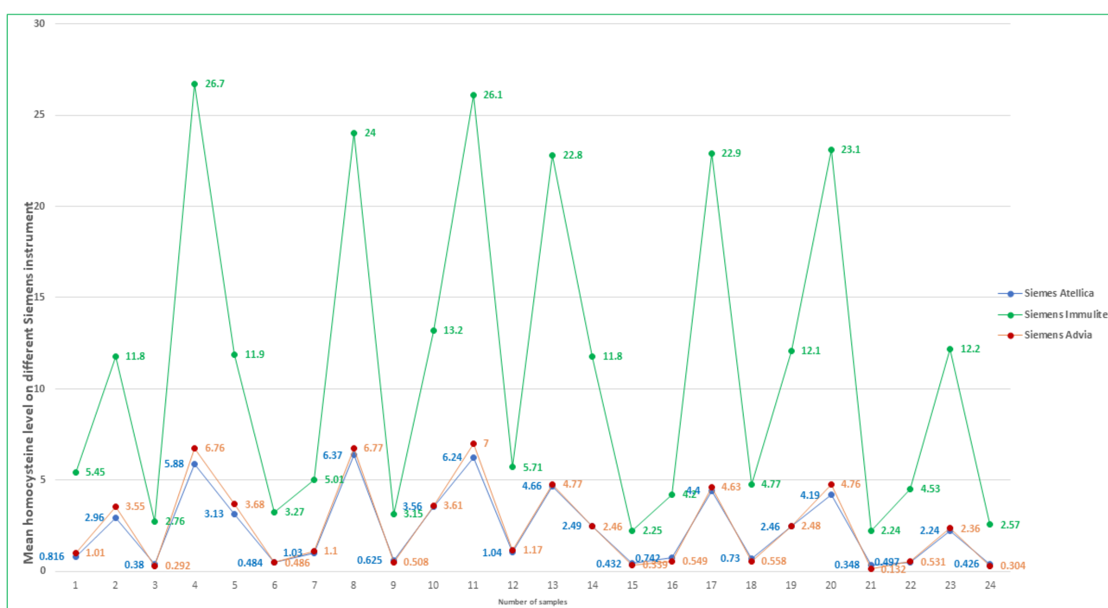
Figure 5: Scatter Plot: Relationship Between Homocysteine Concentration and Bias Percentage.



Scatter plot showing the inverse relationship between homocysteine concentration (x-axis) and bias percentage (y-axis) for the Siemens Advia Centaur XP platform. Peer group bias (blue circles) and method group bias (orange triangles) both demonstrate systematic negative bias across all concentrations, with bias magnitude exceeding −60% at concentrations below 2 µmol/L and persisting at −20% to −30% even at higher concentrations (10–15 µmol/L). This pattern indicates calibrator-related systematic bias rather than random analytical error.

The study also reveals that the difference in evaluation of the level of homocysteine by Siemens Immulite varies widely when compared to Siemens Atellica and Siemens Advia, as illustrated in Figure 6. The figure also shows that the level of homocysteine detected by Siemens Atellica and Siemens Advia is much lower than that of the analyzer of the same manufacturer (Siemens Immulite).

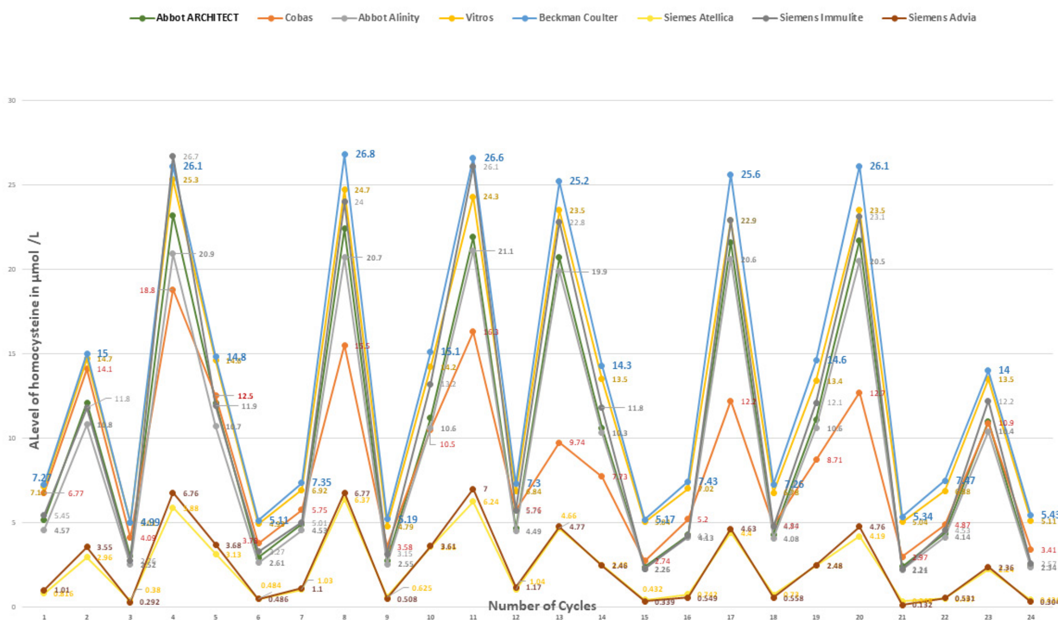
Figure 6: Inter-Platform Comparison Among Siemens Analyzers.



Line graph comparing homocysteine measurements across three Siemens platforms: Immulite (Group 1, high-reading), Advia Centaur XP (Group 3, low-reading), and Atellica (Group 3, low-reading) for all 24 EQA samples. Despite common manufacturer origin, Siemens Immulite consistently reports 30–80% higher homocysteine concentrations compared to Advia and Atellica platforms, highlighting significant intra-manufacturer variability attributable to assay design and calibration strategy differences.

The reported result of the level of homocysteine of 24 samples of External Quality Assurance Services on machines such as Abbot Architect, Cobas, Abbot Alinity, Vitros, Beckman Coulter, Siemens Immulite, Siemens Advia, and Siemens Atellica are shown in Figure 7. The results given by group-1 machines (Abbot Architect, Abbot Alinity, Vitros, Beckman Coulter, SiemensImmulite) are relatively coherent with each other but are much higher than the results declared by group-3 machines (Siemens Advia and Siemens Atellica). The result declared by group-2 machines (Cobas) is sometimes coherent with results declared by group-1 machines, whereas at times, the results declared by it are intermediate between group-1 and group-3 machines.

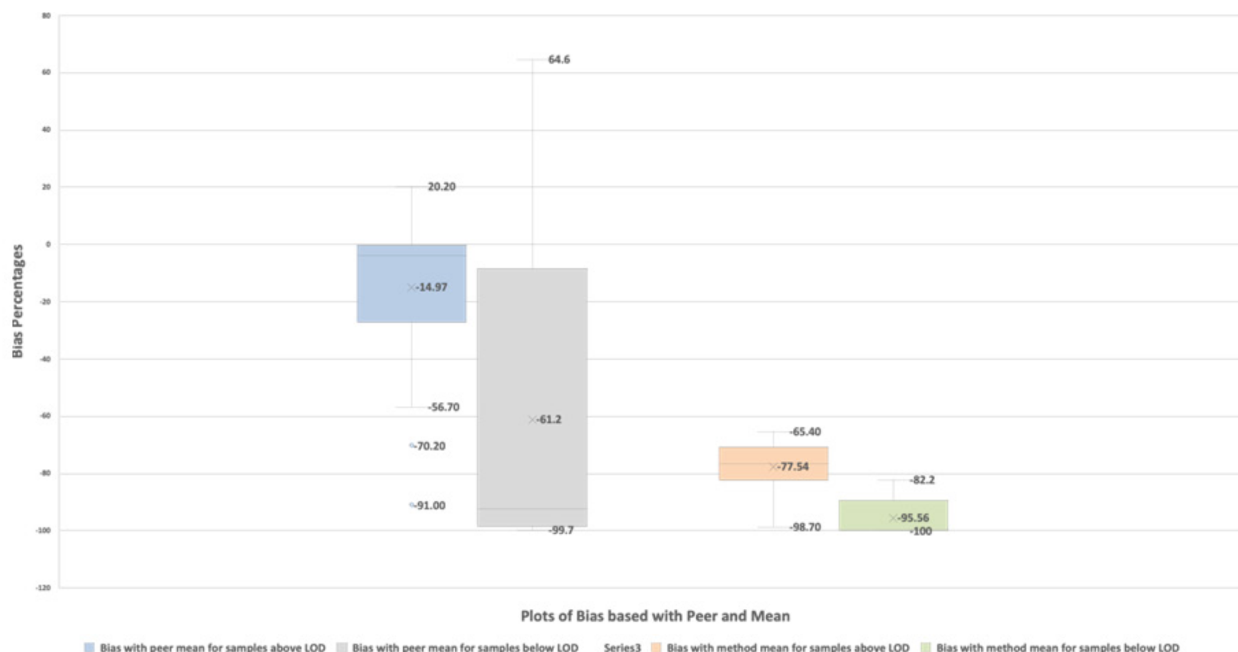
Figure 7: Multi-Platform Comparison: Performance Clustering Analysis.



Multi-line graph illustrating homocysteine measurements across all eight analytical platforms for 24 EQA samples. Three distinct performance clusters emerge: ***Group 1 (high-reading)****: Abbott Architect, Abbott Alinity, Vitros, Beckman Coulter, and Siemens Immulite show concordant results with the highest reported values;* ***Group 2 (intermediate)****: Cobas exhibits variable performance, sometimes aligning with Group 1 and at other times reporting intermediate values;* ***Group 3 (low-reading)****: Siemens Advia and Atellica consistently report the lowest values, systematically diverging from Groups 1 and 2 by 30–70%. Platforms with NIST SRM 1955-traceable calibrators (Cobas, Beckman) demonstrate better concordance with Group 1.*

The box-whisker plot in Figure 8 demonstrates the bias percentage of the lab mean with the peer mean and method mean based on the limit of detection of the Siemens Advia homocysteine kit. Figure 8 shows that the bias percentage of the lab means with peer mean for the values within the limit of detection (blue-colored plot) is less negatively skewed, with a median bias percentage of -14.97%, whereas the plot with values below the limit of detection (grey-colored plot) is more negatively skewed with a median bias percentage of -61.2%. Similarly, the comparison of bias percentage with the method means reveals that the plot is more negatively skewed for the values below the limit of detection (green-colored plot) as compared to the values at or above the limit of detection (orange-colored plot).

Figure 8: Box-Whisker Plot: Impact of LOD on Bias Distribution.



*Box-whisker plots comparing bias percentage distributions for samples within (≥0.5 μmol/L) versus below (<0.5 μmol/L) the Siemens Advia LOD. **Panel A (Peer group bias)**: Samples within LOD show median bias of −14.97% (IQR: −8.2% to −22.1%, blue boxes) with less negative skew, while samples below LOD exhibit median bias of −61.2% (IQR: −48.5% to −75.8%, gray boxes) with marked negative skew and wider dispersion. **Panel B (Method group bias)**: Similar pattern observed with median bias of −12.3% (within LOD, orange boxes) versus −54.7% (below LOD, green boxes). Whiskers represent 1.5×IQR; outliers shown as individual points. This analysis demonstrates that measurements below the analytical LOD dramatically increase bias and imprecision.*

Discussion

The current study involves the analysis of two cycles of Bio-Rad Cardiac Markers Program External Quality Assurance Services Samples (12 each). The Bio-Rad Cardiac Markers External Quality Assurance Services Program is prepared from human serum, plasma, and proteins with added chemicals and preservatives and is accredited as per ISO/IEC 17043:2010 guidelines. Commutability is a characteristic of reference material that determines that, within acceptable limits, the reference material would analytically respond in the same manner as a clinical sample for a measurand evaluated applying various measurement procedures [21-23]. Traceability is the property of reference material that ensures that the results are accurate and are comparable over time and location. It facilitates the global approach based on the preparation, adoption, and use of higher-order international standard reference material and measurement procedures [24]. Bio-Rad External Quality Assurance Services is a peer comparison programme and does not intend to assess the trueness of measurement.

In this study, we have shown that checking the accuracy of clinical parameters such as homocysteine below the detection limit increases the percentage of bias. Interestingly, values below the normal range of human homocysteine levels have hardly any clinical relevance. To the best of our knowledge, there are very few publications on hypohomocysteinemia, whereas there is research on hyperhomocysteinemia leading to cardiovascular problems, neurological and psychiatric problems, and vitamin deficiencies. The current study, for the first time brings out the evident underestimation of homocysteine levels by Siemens Advia and Siemens Atellica platforms when compared with other platforms.

The Standard Reference Material, SRM 1955, from the National Institute of Standards and Technology, is the reference material for Homocysteine and Folate in Human Serum. Homocysteine is analyzed using enzymatic methods on Cobas platforms by a novel enzyme cycling assay. There is a conversion of NADH to NAD+, which is measured at 340 nm. For Cobas platforms, the homocysteine assay kit is calibrated using a 5-point calibration with the homocysteine Calibrator Kit, and the method has been standardized against NIST SRM 1955 reference material. Beckman Coulter also uses an enzymatic method for the estimation of homocysteine, where the conversion of NADH to NAD+ is directly proportional to the concentration of homocysteine. Beckman Coulter also uses calibrators that are prepared gravimetrically and are traceable to NIST SRM 1955, confirmed by a designated measurement procedure (High Performance Liquid Chromatography).

Siemens Immulite uses a competitive immunoassay for the estimation of homocysteine. Homocysteine in the sample is converted to S-adenosyl-homocysteine. The converted S-adenosyl-homocysteine from the patient sample competes with the immobilized S-adenosyl-homocysteine for binding with the alkaline phosphatase labeled-anti-S-adenosyl-homocysteine antibody conjugate. Calibration is performed using Homocysteine Adjustors (LHOL, LHOH), 2 mL each, of synthetically derived S-adenosyl-L-homocysteine in a protein/buffer matrix. Data regarding traceability is not mentioned in the inserts. The Abbot Architect homocysteine assay is a chemiluminescent microparticle immunoassay where the converted S-adenosyl-homocysteine

competes with acridinium-labelled S-adenosyl cysteine for particle-bound monoclonal antibody. The Abbot Architect homocysteine assay uses 1L71-01 Architect Homocysteine Calibrators. Data regarding traceability is not mentioned in the inserts.

Interestingly, the Siemens Advia Centaur homocysteine assay is also a competitive immunoassay using direct, chemiluminescent technology. Converted S-adenosyl-homocysteine from the patient sample competes with S-adenosyl-homocysteine covalently coupled to paramagnetic particles in the Solid Phase for a limited amount of acridinium ester-labeled anti-S-adenosyl-homocysteine in the Lite Reagent. However, the Siemens Advia Centaur homocysteine assay is traceable to an internal standard manufactured using highly purified material. Assigned values of calibrators are traceable to this standardization only. Standardization is achieved when results are equivalent and traceable to a reference measurement procedure [25]. Harmonization is established when results are equivalent, but neither a high-order primary reference material nor a reference measurement procedure is available [26]. The use of Standard Reference Material for traceability of calibrators will ensure the standardization of assays between different manufacturers. External quality assessment materials are not necessarily designed to assess trueness. However, data from an external quality assessment had been shown to aid the efforts in the harmonization of assays [27]. Similarly, the analytical variability in assays on different platforms can also be brought out through the analysis of external quality assessment data [28]. The variation observed in the External Quality Assurance Services results of homocysteine highlights the impending requirement of harmonization of assays using Standard Reference Material. Large-scale data from proficiency testing using traceable samples will yield more conclusive evidence regarding the differences in estimation of analytes on various platforms due to a lack of harmonization. The standardization of different assays would more contribute to the efforts of precise estimation of analytes in patient samples than validating a reference method when different are available for estimation.

Table 1: Analytical Characteristics and Calibrator Traceability of Homocysteine Assays by Platform.

Platform	Manufacturer	Assay Principle	Measuring Range (µmol/L)	LLoQ (µmol/L)	Within-Run Precision (%CV)	Calibrator Traceability
Abbott Architect	Abbott Laboratories	CMIA (competitive, SAH)	1.0 – 50.0	1.0	< 5.0%	Abbott Homocysteine Calibrators
Abbott Alinity	Abbott Laboratories	CMIA (competitive, SAH)	0.90 – 50.0	0.90	< 5.0%	Abbott Homocysteine Calibrators
Roche Cobas	Roche Diagnostics	Enzymatic Cycling	2.0 – 50.0	2.0	< 3.5%	NIST SRM 1955
Ortho Vitros	Ortho Clinical Diagnostics	Enzymatic Colorimetric	2.2 – 45.0	2.2	< 5.0%	In-house prepared calibrator
Beckman Coulter AU / DxC	Beckman Coulter	Enzymatic (NADH detection)	1.5 – 50.0	1.5	< 4.5%	NIST SRM 1955
Siemens Immulite	Siemens Healthineers	Competitive Immunoassay	3.0 – 50.0	3.0	< 7.5%	Siemens Homocysteine Calibrator
Siemens Advia Centaur	Siemens Healthineers	Chemiluminescent Immunoassay	0.5 – 65.0	0.5	< 5.5%	In-house master calibrator
Siemens Atellica	Siemens Healthineers	Chemiluminescent Immunoassay	0.5 – 65.0	0.5	< 5.0%	In-house master calibrator

CMIA, chemiluminescent microparticle immunoassay; SAH, S-adenosylhomocysteine; LOD, limit of detection; NIST, National Institute of Standards and Technology; SRM, Standard Reference Material. Bold indicates confirmed SRM 1955 traceability.

Table 2: Distribution of EQA Samples by Homocysteine Concentration Range.

Concentration Range	Number of Samples (n=24)	Percentage	Clinical Relevance
Below analytical range (<0.5 µmol/L)	5	20.8%	None - below LOD
Hypohomocysteinemia (0.5–4.9 µmol/L)	16	66.7%	Minimal- rarely investigated
Physiological range (5.0–15.0 µmol/L)	3	12.5%	Normal - reference interval
Mild hyperhomocysteinemia (15.1–30.0 µmol/L)	0	0%	High- cardiovascular risk
Moderate–severe hyperhomocysteinemia (>30.0 µmol/L)	0	0%	Very high -requires treatment

Based on peer group mean values from 24 Bio-Rad EQA samples (Cycles 10 and 11).

Conclusion

In conclusion, our study demonstrates the wide difference in the bias percentage of homocysteine levels when bias was calculated with the peer mean and the method mean. There was a substantial increase in bias percentage for the Siemens Advia Centaur XP machine for samples with levels of homocysteine below the limit of detection. Our analysis also underscores the importance of standardization in the clinical analysis of homocysteine levels. While elevated homocysteine is well-documented in contributing to cardiovascular, neurological, and psychiatric pathologies, the clinical significance of low homocysteine levels remains unsubstantiated. Ensuring consistency in homocysteine measurement is essential for reliable diagnosis and effective treatment of related diseases. We suggest the standardization of assays for the evaluation of homocysteine with the use of Standard Reference Material in order to ensure consistency in evaluation of homocysteine levels using various platforms across the country.

Acknowledgments

The authors acknowledge the support of the laboratory staff and technical personnel who contributed to this study. We thank the Bio-Rad External Quality Assurance Services program for providing proficiency testing services.

Declaration of conflict of interests

There is no competing financial interest or personal relationship that could have appeared to influence the work reported in this paper.

Ethical approval

This study involved analysis of de-identified proficiency testing data without access to patient information or samples. No ethical approval was required for analysis of anonymized quality control data according to institutional policies. All studies involving human subjects are in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Credit author statement

Sojit Tomo: Conceptualization, Investigation, Formal analysis, Writing - Original Draft, Visualization Snigdha Singh: Investigation, Data Curation, Writing - Review & Editing Amandeep Birdi: Methodology, Validation, Writing - Review & Editing Manoj Khokhar: Resources, Data Curation, Writing - Review & Editing Dharmveer Yadav: Formal analysis, Visualization, Writing Review & Editing Mithu Banerjee: Conceptualization, Methodology, Supervision, Project administration, Writing - Review & Editing

Funding statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability statement

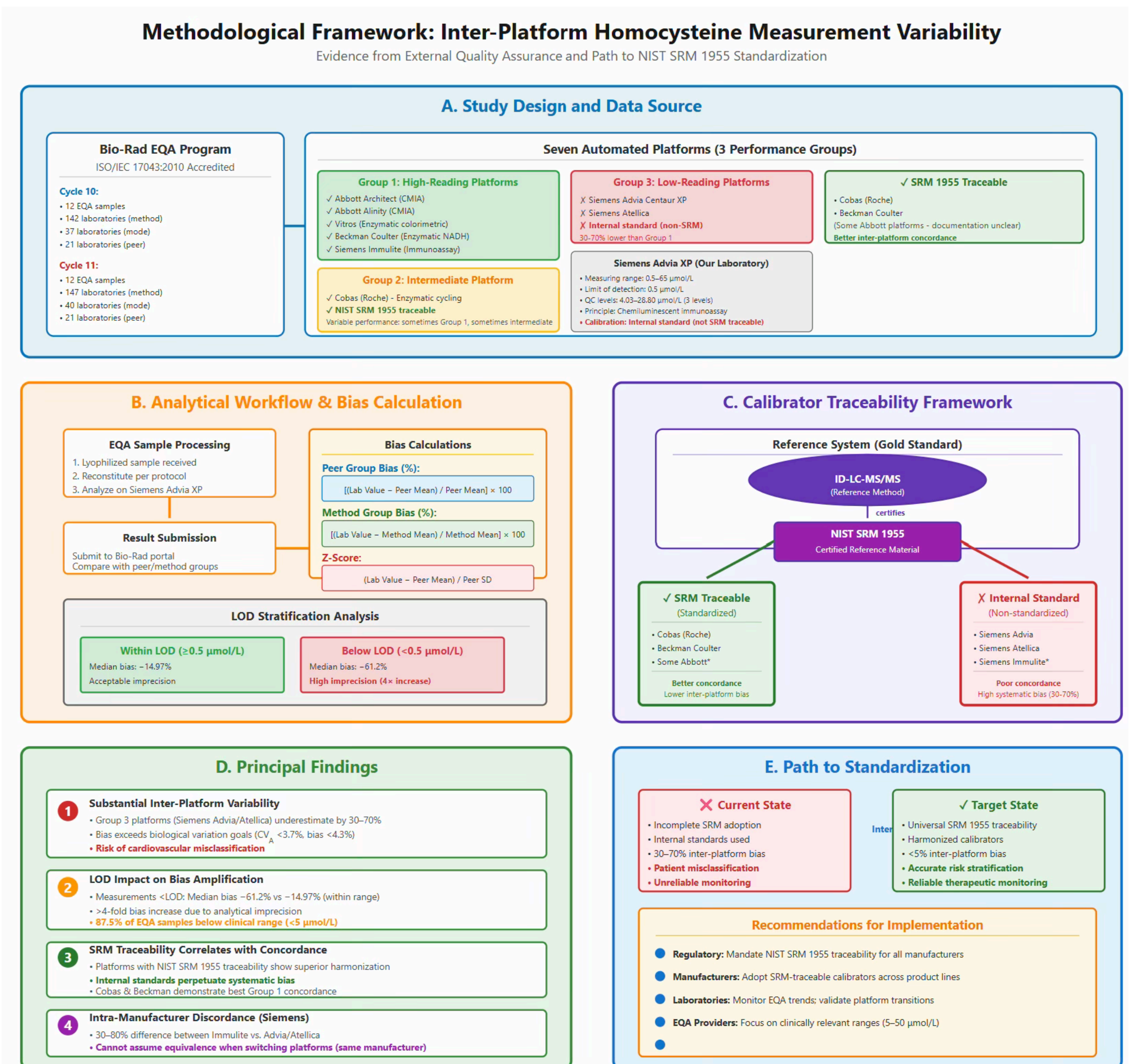
The data supporting the findings of this study are available from the Bio-Rad External Quality Assurance Services program reports. Data are available from the authors upon reasonable request and with permission of Bio-Rad Laboratories.

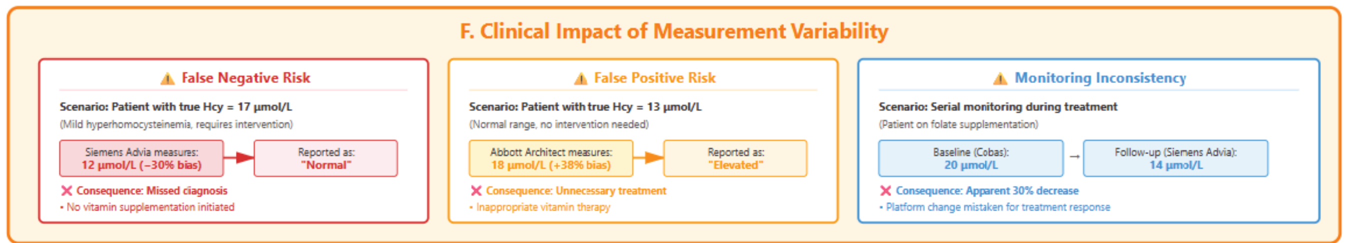
References

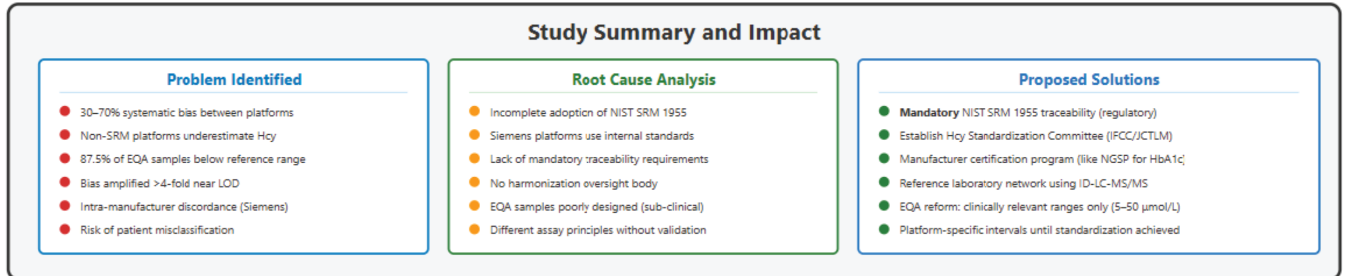
1. Nelson BC, Pfeiffer CM, Zhang M, Duewer DL, Sharpless KE, Lippa KA. Commutability of NIST SRM 1955 Homocysteine and Folate in Frozen Human Serum with selected total homocysteine immunoassays and enzymatic assays. Clin Chim Acta 2008;395:99-105. https://doi.org/10.1016/j.cca.2008.05.012
2. Varga EA, Sturm AC, Misita CP, Moll S. Homocysteine and MTHFR Mutations. Circulation 2005;111:e289-e293. https://doi.org/10.1161/01.CIR.0000165142.37711.E7
3. Acharya U, Gau JT, Horvath W, Ventura P, Hsueh CT, Carlsen W. Hemolysis and hyperhomocysteinemia caused by cobalamin deficiency: three case reports and review of the literature. J Hematol Oncol 2008;1:26. https://doi.org/10.1186/1756-8722-1-26

4. Ding C, Li J, Wei Y, et al. Associations of total homocysteine and kidney function with all-cause and cause-specific mortality in hypertensive patients: a mediation and joint analysis. HypertensRes 2024;47:1500-1511. https://doi.org/10.1038/s41440-024-01615-8
5. Nygård O, Refsum H, Ueland PM, Vollset SE. Major lifestyle determinants of plasma total homocysteine distribution: the Hordaland Homocysteine Study. Am J Clin Nutr 1998;67:263-270. https://doi.org/10.1093/ajcn/67.2.263
6. de Bree A, Verschuren WM, Blom HJ, Kromhout D. Lifestyle factors and plasma homocysteine concentrations in a general population sample. Am J Epidemiol 2001;154:150-154. https://doi.org/10.1093/aje/154.2.150
7. Desouza C, Keebler M, McNamara DB, Fonseca V. Drugs affecting homocysteine metabolism: impact on cardiovascular risk. Drugs 2002;62:605-616. https://doi.org/10.2165/00003495-200262040-00003
8. Finkelstein JD, Martin JJ. Homocysteine. Int J Biochem Cell Biol 2000;32:385-389. https://doi.org/10.1016/s1357-2725(99)00138-7
9. Refsum H, Ueland PM, Nygård O, Vollset SE. Homocysteine and cardiovascular disease. Annu Rev Med 1998;49:31-62. https://doi.org/10.1146/annurev.med.49.1.31
10. Lee R, Frenkel EP. Hyperhomocysteinemia and thrombosis. Hematol Oncol Clin North Am 2003;17:85-102. https://doi.org/10.1016/s0889-8588(02)00069-4
11. Hermann A, Sitdikova G. Homocysteine: Biochemistry, Molecular Biology and Role in Disease. Biomolecules 2021;11:737. https://doi.org/10.3390/biom11050737
12. Loscalzo J. The oxidant stress of hyperhomocyst(e)inemia. J Clin Invest 1996;98:5-7. https://doi.org/10.1172/JCI118776
13. Kaplan P, Tatarkova Z, Sivonova MK, Racay P, Lehotsky J. Homocysteine and Mitochondria in Cardiovascular and Cerebrovascular Systems. Int J Mol Sci 2020;21:7698. https://doi.org/10.3390/ijms21207698
14. Yakovleva O, Bogatova K, Mukhtarova R, et al. Hydrogen Sulfide Alleviates Anxiety, Motor, and Cognitive Dysfunctions in Rats with Maternal Hyperhomocysteinemia via Mitigation of Oxidative Stress. Biomolecules 2020;10:995. https://doi.org/10.3390/biom10070995
15. Welch GN, Loscalzo J. Homocysteine and atherothrombosis. N Engl J Med 1998;338:1042-1050. https://doi.org/10.1056/NEJM199804093381507
16. Elsherbiny NM, Sharma I, Kira D, et al. Homocysteine Induces Inflammation in Retina and Brain. Biomolecules 2020;10:393. https://doi.org/10.3390/biom10030393
17. Boushey CJ, Beresford SA, Omenn GS, Motulsky AG. A quantitative assessment of plasma homocysteine as a risk factor for vascular disease. Probable benefits of increasing folic acid intakes. JAMA 1995;274:1049-1057. https://doi.org/10.1001/jama.1995.03530130055028
18. Karava V, Kondou A, Dotis J, et al. Hemolytic Uremic Syndrome Due to Methylmalonic Acidemia and Homocystinuria in an Infant: A Case Report and Literature Review. Children (Basel) 2021;8:112. https://doi.org/10.3390/children8020112
19. Caldeira-Araújo H, Ramos R, Florindo C, Rivera I, Castro R, Tavares de Almeida I. Homocysteine Metabolism in Children and Adolescents: Influence of Age on Plasma Biomarkers and Correspondent Genotype Interactions. Nutrients 2019;11:646. https://doi.org/10.3390/nu11030646
20. Azzini E, Ruggeri S, Polito A. Homocysteine: Its Possible Emerging Role in At-Risk Population Groups. Int J Mol Sci 2020;21:1421. https://doi.org/10.3390/ijms21041421
21. Badrick T, Punyalack W, Graham P. Commutability and traceability in EQA programs. Clin Biochem 2018;56:102-104. https://doi.org/10.1016/j.clinbiochem.2018.04.004
22. Xing T, Liu J, Sun H, Gao Y, Ju Y, Liu X, Song D. Commutability assessment of reference materials for homocysteine. Clin Chem Lab Med 2022;60:1562-1569. https://doi.org/10.1515/cclm-2022-0594
23. Vogeser M, Habler K. Is commutability of a reference material always desirable? J Mass Spectrom Adv Clin Lab 2023;31:17-18. https://doi.org/10.1016/j.jmsacl.2023.11.001
24. Beastall GH. Traceability in Laboratory Medicine: What is it and Why is it Important for Patients? EJIFCC 2018;29:242-247.
25. Plebani M. Harmonization in laboratory medicine: Requests, samples, measurements and reports. Crit Rev Clin Lab Sci 2016;53:184-196. https://doi.org/10.3109/10408363.2015.1116851
26. Plebani M. Harmonization in laboratory medicine: the complete picture. Clin Chem Lab Med 2013;51:741-751. https://doi.org/10.1515/cclm-2012-0812
27. Clerico A, Ripoli A, Zucchelli GC, Plebani M. Harmonization protocols for thyroid stimulating hormone (TSH) immunoassays: different approaches based on the consensus mean value. Clin Chem Lab Med 2015;53:377-382. https://doi.org/10.1515/cclm-2014-0586
28. Jassam N, Weykamp C, Thomas A, et al. Post-standardization of routine creatinine assays: are they suitable for clinical applications. Ann Clin Biochem 2017;54:386-394. https://doi.org/10.1177/0004563216661961

Supplementary Figure 1: Comprehensive methodological framework illustrating inter-platform variability in homocysteine measurement, calibrator traceability pathways, and clinical implications of measurement bias.







Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

National Survey on Delta check Practices in Clinical Laboratories Across Pakistan

Fatima Kanani[1*], Saba Raza[1], Harim Fatima[1], Adnan Mustafa Zubairi[1]

[1]Section of Chemical Pathology, Indus Hospital and Health Network, Karachi, Pakistan

Article Info

Corresponding Author:

Fatima Kanani

Section of Chemical Pathology

Indus Hospital and Health Network, Karachi, Pakistan

Phone: 0092300 2102283

E-mail: fatima.kanani@tih.org.pk

Keywords

Delta check, Patient-based real time quality control, Standardization, Survey, Delta alerts

Abstract

Background: Delta check is a quality control tool in clinical laboratory that compares a patient's current test results with previous ones. It is often considered to be a useful tool to identify pre-analytical errors. Delta check procedures vary across labs. A survey was conducted to determine national delta check practices.

Method: An online survey was sent to pathologists. Data was analyzed by Microsoft excel and figures were generated via R statistical software.

Results: The survey was sent out to 84 participants. There were 30 (35.7%) final responses, after exclusion of duplicates and incomplete entries. Respondents varied in their designations from Professor 2(6.7%) to Assistant Professor 10(33.3%). Most respondents 25 (83.3%) were Fellows of College of Physicians and Surgeons, while there were 3(10%) PhDs. Work experience varied from less than 3 years, 5(16.7%) to more than 10 years, 17(56.7%). 25 labs surveyed were from tertiary care hospitals. Most laboratories were in Sindh Province, 13 (43.3%). Delta check procedures were used by 19(63.3%) labs, of which 6(31.6%) used automated checks, 12(63.2%) had defined delta limits for different tests, 15(78.9%) had defined time limits to check delta results, 17(89.5%) had a checklist for its investigation and 10(52.6%) had a Laboratory Information System or middleware to block failed delta results. Most of the respondents considered the smallest available option as the ideal delta difference value for blocking results.

Conclusion: Delta check procedures vary among surveyed labs and efforts are required for the formulation of national guidelines to synchronise and standardize delta check practices.

Introduction

Delta check is a mode of patient-based real time QC (PBRTQC) used in clinical laboratories in which difference between sequential patient results help to identify potential pre analytical, analytical or post analytical errors. The earliest work on delta check procedures in clinical laboratory is from 1970s [1, 2]. Its greatest utility is suggested to be in picking up pre-analytical errors [3-5], although the use of wrist bands for patients' identification and scanners for its confirmation, bedside bar code labelling, electronic ordering and receiving of specimens, and automated hemolysis, icterus, lipemia (HIL) checks have reduced these errors [6]. With the advent of automation, interfacing of instruments with Laboratory Information System (LIS) and use of middle wares, post analytical errors such as transcriptional or calculation mistakes have also decreased. Analytical faults account for the least percentage of these errors. The most common cause of significant difference between two serial results is due to change in physiological condition [5, 7].
Delta checks are also an inherent part of most auto-verification algorithms and have been recommended in various latest guidelines[8-10]. Clinical and Laboratory Standards Institute (CLSI) guideline EP33 ED2:2023 explains in depth the calculations and utilities of delta checks [11], while College of American Pathologists (CAP) has incorporated clauses related to delta checks in the accreditation checklist's section on auto-verification.
National surveys are a good tool for gaining an insight into the current practices and identifying the need to improve upon them. Data collection is relatively simple. A drawback is the reliance on the participants for the availability and validity of the responses. The aim of conducting this survey was to acquire baseline knowledge regarding current delta check procedures across different laboratories in the country and to identify the existing gaps. The results of the survey will be used for capacity building of the clinical laboratory fraternity, and aim to make recommendations for formulation of national guidelines for standardizing the use of delta checks in reporting patient test results.

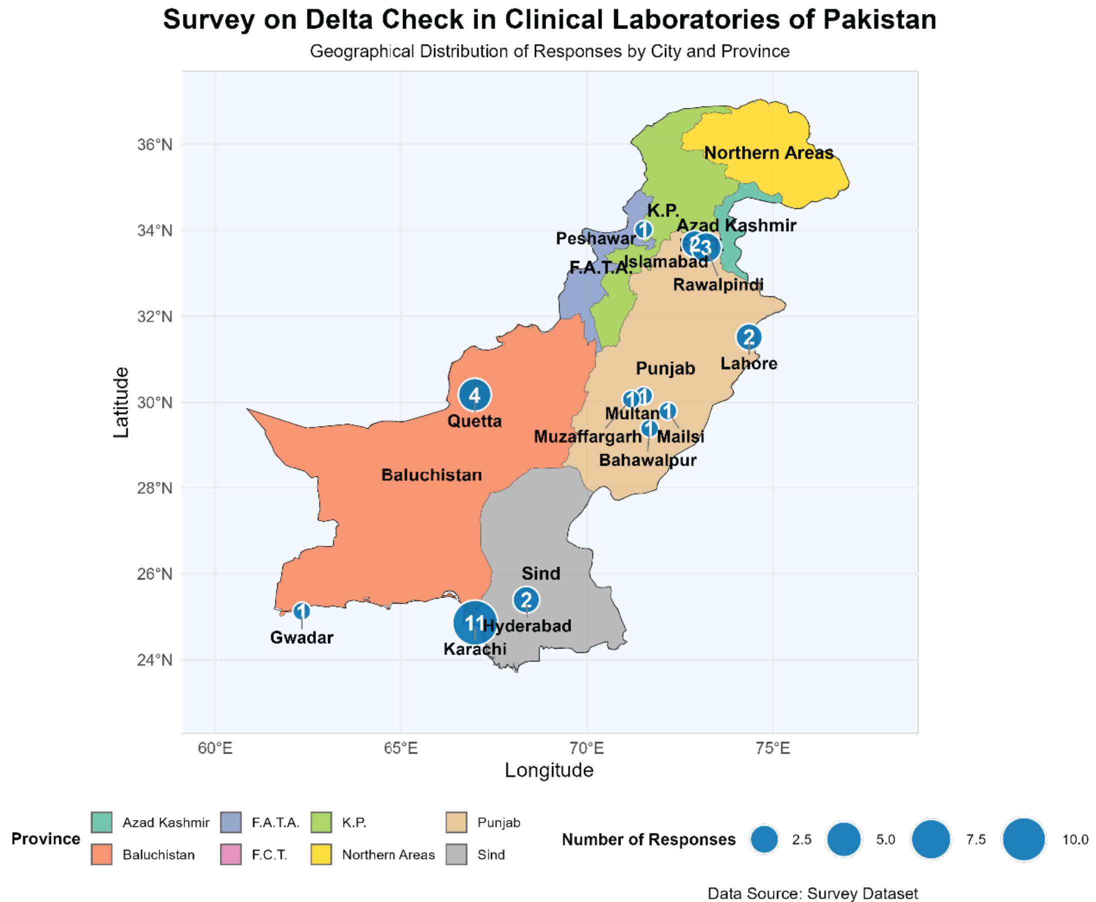
Material and Methods

The survey was conducted from December 2024 to January 2025 after IRB approval # IHHN_IRB_2024_11_004 and in accordance with Declaration of Helsinki (2024). The questionnaire was prepared on RedCap which is a web-based questionnaire tool, and sent out to a total of 84 pathologists working in clinical laboratories across Pakistan selected by snowball technique. Pathologists with whom contact already existed were asked to give contacts of other potential participants. Consent of the participants was taken in the questionnaire. Inclusion criteria included Chemical Pathologists, Haematologists, Heads of department or consultants working in clinical laboratories across Pakistan and holding Fellowship of College of Physicians and Surgeons, Pakistan (FCPS), Masters of Philosophy (MPhil) or Doctor of Philosophy (PhD) degrees. Exclusion criteria included duplicate entries from same lab and incomplete responses. The questionnaire was sent via WhatsApp or email to the various participants. In the first section of the survey, general details of the participants and the hospital/ laboratory was gathered. The second part dealt with their laboratory practices related to delta check procedures. The participants were then asked about what delta check limits they thought should be used for various laboratory tests, including calcium, sodium, potassium, chloride, bicarbonate (HCO3), urea, creatinine, uric acid, cardiac troponins, HbA1c, haemoglobin, white blood cell (WBC) count, platelet count and International Normalized ratio (INR).
Results were analyzed by Microsoft Excel version 2511. Frequencies and percentages were calculated for the various variables. R statistical software was used to create figures.

Results

The survey was sent out to 84 participants. There were 30 (35.7%) final responses, after exclusion of 4 duplicates and one incomplete entry. Respondents varied in their designations from Professor 2 (6.7%) to Assistant Professor, 10 (33.3%). Majority of the respondents 25 (83.3%) were FCPS, while there were 3 (10%) PhDs. Work experience varied from less than three years, 5 (16.7%) to more than 10 years, 17 (56.7%), while there were 4 (13.3%) pathologists each with 3-5 and 5-10 years of experience. The survey participants included pathologists from Federal Capital Territory and all four provinces of Pakistan, (Sindh, Baluchistan, Punjab and Khaybar Pakhtunkhwa). Sindh had the largest representation, 13 (43.3%) of which the city of Karachi had the most labs, 11 (36.6%). Next highest representation was from Punjab 9 (30%), (Figure 1).

Figure 1: Map of Pakistan showing number of labs and their cities included in the survey.



The respondents belonged to various laboratory types, predominantly from labs based in tertiary care hospitals, 25 (83.3%), while 4 (13.3%) were from labs in secondary care hospitals and 1 (3.3%) from a stand-alone laboratory. Hospitals of various bed strengths were covered and majority of them were moderately sized having 100-500 beds, 19 (65.5%). Survey questions and the responses regarding the clinical laboratories' practices are mentioned in Table 2.

Table 2: Responses to questions related to laboratory practices.

Variable	Options	Frequency (Percentage)
Laboratory type (n=30)	Laboratory in a secondary care hospital	4 (13.3)
	Laboratory in a tertiary care hospital	25 (83.3)
	Stand-alone laboratory	1 (3.3)
Hospital beds (n=29)	0-50	1 (3.4)
	50-100	1 (3.4)
	100-500	19 (65.5)
	> 500	8 (27.6)
Clinical laboratory accreditation	No	12 (40)
	Yes	18 (60)
If yes: (n=18)	CAP-accredited	3 (16.7)
	ISO 15189 accredited	8 (44.4)
	Other	7 (38.9)
Presence of laboratory information system in clinical laboratory (n=30)	No	6 (20)
	Yes	24 (80)

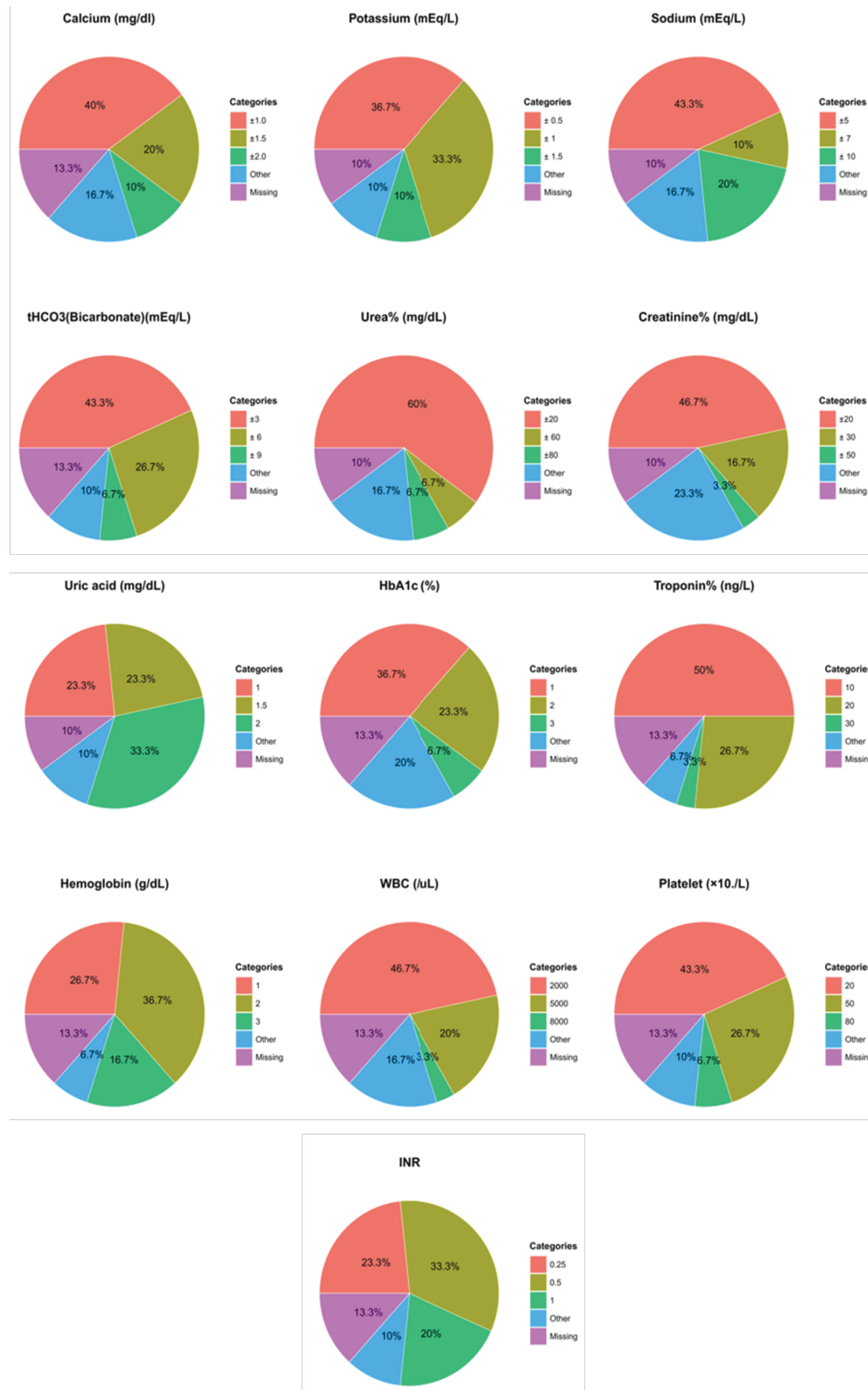
Variable	Options	Frequency (Percentage)
Mode of test reports in clinical laboratory? (n=30)	Electronic reports	6 (20)
	Printed paper reports	4 (13.3)
	Both electronic and printed reports	20 (66.7)
Interfacing of results to laboratory information system and reports (n=30)	No	10 (33.3)
	Yes	20 (66.7)
Auto-verification procedures (n=30)	No	22 (73.3)
	Yes	8 (26.7)
Delta check procedure before releasing the patients results (n=30)	No	11 (36.7)
	Yes	19 (63.3)
Delta check procedure manual or automated (n=19)	Automated	6 (31.6)
	Manual	13 (68.4)
Time period defined for checking delta results (n=19)	No	4 (21.0)
	Yes	15 (78.9)
If yes, what is the defined time period? (n=15)	1 day	1 (6.7)
	1 week	1 (6.7)
	3 days	2 (13.3)
	Dependent on the type of the test	11 (73.3)
Are parameters defined for which a delta check is required? (n=19)	No	9 (47.4)
	Yes	10 (52.6)
Are Delta check limits defined for each type of test or dependent on the laboratory staff? (n=19)	No, dependent on the laboratory staff	7 (36.8)
	Yes, defined delta check limit	12 (63.2)
How are delta limits defined? (n=12)	From the literature survey	6 (50)
	From the labs own data	4 (33.3)
	By Reference Change Value (RCV) based on intra/ inter-individual variation	2 (16.7)
Checklist to work out the possible causes of delta failure (n=19)	No	2 (10.5)
	Yes	17 (89.5)

CAP: College of American Pathologists, ISO: International Organization for Standardization, LIS: Laboratory Information System

There were 18 clinical laboratories which were accredited out of which 3 (16.7%) labs were accredited by CAP, while 8 (44.4%) were International Organization for Standardization (ISO) 15189 accredited. Other labs mentioned accreditation by ISO 9000, ISO 9001, Punjab Healthcare Commission and Joint Commission International (JCIA). Most of the labs surveyed, 24 (80%) had an LIS and 20 (66.7%) percent said that their laboratory results were interfaced to LIS. A total of 8 labs (26.7%) claimed to have an auto- verification system, while 19 (63.3%) said they have placed a procedure for delta checks before releasing patient results. The specific variations in delta check practices are mentioned in Table 2.

The opinion of the participants for delta check limits of various common analytes is shown in Figure 2. When a participant did not give any opinion for these delta check limits, it is labelled as missing. In the opinion of 11 (37%) participants appropriate delta for potassium was ±0.5mEq/L while delta limit for sodium was suggested as ±5 mEq/L by 13(43%) respondents. For both urea, 18 (60%) and creatinine, 14 (47%) participants considered ±20% change as the most suitable delta limit.

Figure 2: Participants' opinion on optimal delta limits for common analytes.



Discussion

This is the first survey conducted in Pakistan on delta checks practices in clinical laboratories. Amongst the labs surveyed, 19 (63.3%) used delta check procedures in their labs. A lab survey conducted on clinical laboratories in Spain revealed a response rate of 13.2%, [12], while in China a similar survey showed the response rate as 24.7% [7]. Our response rate in comparison is 30(35.7%). The overall number of labs in our survey is not very high, yet this is the usual response rate received in various clinical laboratory surveys conducted in Pakistan [13, 14]. The higher participation of senior pathologists from medium to large sized labs could be due to their greater understanding and sensitization for the topic. The snowballing technique used in reaching out to the participants had its disadvantages as the immediate contacts responded more frequently than those who were approached through secondary contacts. Hesitation to share information regarding institutional practices, not being familiar or acquainted with the survey topics and lack of time or interest could be the other reasons for not participating in the survey. Since systematic implementation of delta check procedure

requires an advanced LIS or middleware which is not available in most of the clinical laboratories in Pakistan, it is possible that the pathologists without access to these resources refrained from participating in the survey.

A Croatian survey on medical laboratories showed that 30% used delta checks in their labs [15]. In the Spanish survey delta checks were being used in 24% of the labs [12] and in China, delta check was used by 25% of the labs surveyed [7]. Usage of delta check procedures in our survey, 19 (63.3%), is probably greater compared to other surveys because of more participation by tertiary care institutes in this survey.

The analytes considered good candidates for delta are those with low index of individuality or low intra individual variation [4, 16, 17]. Reference change value (RCV), which is based on intra individual biological variation, is accepted by CLSI as a tool to set the limits [11]. In our survey only 2 labs mentioned that they used RCV to set their delta limits, while 6 used literature-based values. However, RCV limits need to be verified and optimized to prevent high number of false alarms [18].

Auto-verification was stated to be present in 8 (26.7%) of the labs surveyed, although the extent of auto-verification was not asked. A global survey published in 2024 covering 920 labs from 55 countries, showed auto-verification was used in 10% labs [9]. Another study showed less than 20% of hospital labs were engaged in auto verification [19]. Overall, as other surveys have reported across Europe and Middle East, currently information technology in clinical laboratory is underutilized, even for the basic functions [9, 19, 20]. An alarming revelation during our survey, which is a testimony to this fact, is that one third of the labs surveyed did not have result interfacing. Implementing a delta check rule, albeit manual, would be of paramount importance in these cases to catch transcription errors.

Out of the 19 labs which had a delta check procedure in place in our survey, majority, 17, had a checklist for working on the delta check failures. Investigation check list or algorithms for trouble shooting delta failures are important to delineate the root cause, and these should focus on all three testing phases.

Ideal time gap between the two results to elicit delta failure can vary depending on the parameter, patient setting (inpatient versus outpatient), and other factors. This varies greatly with the laboratory practice worldwide [4]. Clinical specimens are not generally stored for more than one week and there are stability issues for many parameters. Previous specimen is also ideally required for delta investigation. These factors together call for practical constraints in setting delta time limits. Keeping a long time period may give more false alerts as certain analytes may change over time for variable reasons, for example; disease state, patient management, transfusion and dialysis. An exception is HbA1c in which the optimal time is three months because of red blood cell turnover time. Keeping variable time frames for different tests might be hindered by the available LIS or middleware in the lab. CAP Q probe study showed that median time span was 3-7 days for the routine assays [5] while Tan et al state that there is no ideal time limit cut off [21]. HbA1c was not included in these studies.

Most of the respondents chose the smallest option as the optimal delta difference for the analytes surveyed. Notable exception was for haemoglobin where majority selected a difference of 2.0 gram/decilitre as a significant change, which was also the median delta limit pointed out in CAP Q probes [5]. For sodium, majority, 13 labs preferred the delta limit as 5.0 mEq/L, while in CAP Q probes study the median sodium delta change limit was 9.0 mEq/L [5]. Similarly for Calcium the delta limit chosen by 12 labs was 1.0 mg/dL, while in the CAP Q probe study a median of 2.0 mg/dL was identified [5]. Another obvious discrepancy was in total CO2 or bicarbonate in which 13 (43%) labs in our survey mentioned that 3.0 mEq/L should be kept as the delta limit, while in CAP Q probes the median delta limit was 8.0 mEq/L [5]. The underlying point is that delta checks limits should be optimized such that there are minimum false alarms. Physiological changes account for most delta alerts. A very narrow range can give unnecessary alarms thus increasing the turnaround time and resource wastage. It is important to optimize the delta limits and review them periodically. Similarly, the limits should not be so relaxed as to miss the alarms [4, 17].

Although 19 laboratories reported having a delta check process in place, more responses were received when participants were asked to specify ideal delta limits for various parameters. This suggests that while pathologists had parameter specific delta cut-off values in their minds, these had not been implemented, possibly due to limitations in IT support systems.

The various calculations suggested to determine the delta threshold are absolute delta check, percentage delta check, absolute rate difference delta and percentage rate difference delta. These depend on whether absolute numerical values or percentages are used and with or without time scale. While absolute delta checks are appropriate for parameters such as sodium and chloride where the analytes are kept in strict physiological limits, percentage change is more relevant for urea and creatinine specially at higher values. For an analyte such as neonatal bilirubin, time scale is very significant. Multivariate delta or delta difference across multiple parameters can be more useful to point towards a cause, for example, urea and creatinine, ALT and AST, serum protein and albumin [10]. Gender specific absolute delta thresholds for cardiac troponins have been suggested [22], with researchers suggesting a delta limit of 10 ng/L for men and 7.0 ng/L for women. Urea and creatinine delta limits in patients on dialysis do not hold much value and give false alarms. Other calculations have been proposed for these patients [23]. International Normalized Ratio (INR) is a well-known marker performed widely for screening patients for coagulopathy. Various INR cut-offs have been used for different conditions [24], and the laboratory practices must be aligned with the current clinical practice guidelines. Evidence-based approaches, including retrospective data analysis and receiver operating curve (ROC) evaluation, are increasingly used to set and verify delta check limits tailored to local clinical and operational needs [23, 25, 26]. Multianalyte and logistic regression-based delta checks improve detection of contamination by intravenous fluid and preanalytical errors, reducing false alert rates and enhancing specificity [26, 27]. A clinically pertinent use of delta is a real time alert system based on delta differences for selected clinical conditions, including acute myocardial infarction (AMI) and acute kidney injury

(AKI). In 2012, researchers stated that introducing electronic alerts for AKI could improve short term clinical outcome and intervention [28]. N Flynn and A Dawnay did a study on real-time automated delta check electronic alerts flagging a 50% increase in creatinine to a concentration of >50 mmol/L from the last result within 90 days and saw that 70% of the delta alerts were due to AKI [29]. Consensus document from Acute Disease Quality Initiative meeting concluded that digital health solutions provide enhanced tools to improve AKI care with on going improvements in diagnosis and monitoring [30]. This survey does not give a snapshot of all laboratories in Pakistan, though it does provide a very true picture of the specialized labs and their practices related to delta checks. Detailed information about the methodology and criteria for application of delta rule for each parameter was not asked in the survey. The next phase will be to conduct educational sessions for implementation of delta checks focussing on causes, delta work up checklists, algorithms and calculations. Sporadic educational activities in Pakistan on this topic have taken place over the years, however, with increased awareness and growing number of laboratories, this topic needs to be re-addressed in detail. Consistent usage of optimal delta check for identifying total testing process errors is required. It can also be made a key performance indicator (KPI) for continual quality improvement. Individual variations in delta procedures will exist depending on the laboratory set-up and its requirements. We propose a National Working Group under the umbrella of the Pakistan Society of Chemical Pathology (PSCP) for quality improvement initiatives, with one of the aims to elevate the reporting standards and standardize patient test reporting practices.

Limitations

The number of laboratories surveyed is small, nevertheless, they span Federal Capital and all four provinces, and include a large number of subject experts and some of the major hospitals of the country. Participation and response biases are other limitations of such survey designs [31].

Conclusion

Delta check practices vary across clinical laboratories in Pakistan. There is need to formulate and implement National guidelines for the standardized use of delta check practices in routine clinical laboratory operations.

Conflict of Interest

None declared.

Ethical Statement

Expedited ethical approval was granted by the Institutional review board (IHHN_IRB_2024_11_004) in accordance with Declaration of Helsinki.

Author contributions

Fatima Kanani: Conceptualization, methodology, software, analysis, initial draft, review and editing
Saba Raza: Methodology, review
Harim Fatima: Software, initial draft
Adnan Mustafa Zubairi: Methodology, resources, review

Acknowledgement

We would like to acknowledge Syed Wajahat Hussain for his help in analysis.

Funding

The authors did not receive any funding for this study.

Data Availability

Study data is available on request from the corresponding author according to institutional policies.

References

1. Ladenson JHJCc. Patients as their own controls: use of the computer to identify" laboratory error". Clin Chem. 1975;21(11):1648-1653. DOI:10.1093/clinchem/21.11.1648.
2. Nosanchuk JS, Gottmann AW. CUMS and delta checks. A systematic approach to quality control. American journal of clinical pathology. 1974;62(5):707-712. DOI:10.1093/ajcp/62.5.707.
3. Castro-Castro MJ, Sanchez-Navarro L. Estimation of change limits (deltacheck) in clinical laboratory. Advances in laboratory medicine. 2021;2(3):417-431. DOI:10.1515/almed-2020-0114.
4. Markus C, Tan RZ, Loh TP. Evidence-based approach to setting delta check rules. Crit Rev Clin Lab Sci. 2021;58(1):49-59. DOI:10.1080/10408363.2020.1800585.
5. Schifman RB, Talbert M, Souers RJ. Delta Check Practices and Outcomes: A Q-Probes Study Involving 49 Health Care Facilities and 6541 Delta Check Alerts. Archives of pathology & laboratory medicine. 2017;141(6):813-823. DOI:10.5858/arpa.2016-0161-CP.
6. Loh TP, Tan RZ, Sethi SK, Lim CY, Markus C. Delta checks. Advances in clinical chemistry. 2023;115:175-203. DOI:10.1016/bs.acc.2023.03.005.
7. He S, Kang F, Wang W, Chen B, Wang Z. National survey on delta checks in clinical laboratories in China. Clinical chemistry and laboratory medicine. 2020;58(4):569-576. DOI:10.1515/cclm-2019-1131.
8. Rimac V, Vlasic Tanaskovic J, Jokic A, Honovic L, Podolar S, Lenicek Krleza J. National recommendations of the Working Group for Post-analytics of the Croatian Society of Medical Biochemistry and Laboratory Medicine: implementation of autovalidation procedures. Biochemia medica. 2025;35(1):01-13. DOI:10.11613/BM.2025.010503.
9. Huf W, Mohns M, Almeta E, Lister R, Buchta C, Demyanets S, et al. Benchmarking medical laboratory performance on a global scale. Frontiers in public health. 2024;12:1363957. DOI:10.3389/fpubh.2024.1363957.
10. Randell EW, Yenice S. Delta Checks in the clinical laboratory. Crit Rev Clin Lab Sci. 2019;56(2):75-97. DOI:10.1080/10408363.2018.1540536.
11. Institute CaLS. Use of Delta Checks in the Medical Laboratory, 2nd Edition. Wayne, Pennsylvania, USA: Clinical and Laboratory Standards Institute; 2023 [updated 16-12-25; cited 2025 15-Nov-2025]. Available from: https://clsi.org/shop/standards/ep33/.

12. Gomez-Rioja R, Alvarez V, Ventura M, Alsina MJ, Barba N, Cortes M, et al. Current status of verification practices in clinical biochemistry in Spain. Clinical chemistry and laboratory medicine. 2013;51(9):1739-1746. DOI:10.1515/cclm-2012-0659.
13. Afzal N, Batool H, Raza S, Ayub S, Bashir S, Hayat A, et al. A National e-Survey of Adult Reference Intervals of Routine Chemistry Analytes Used by Laboratories Across Pakistan: A Step Towards Harmonization. EJIFCC. 2025;36(2):132-142.
14. Ahmed S, Khan FMA, Abbas G, Iqbal S, Shafi M, Arbab K, et al. Standardizing The Biochemical Tests for Chronic Kidney Disease (CKD): Where Do We Stand? A National Survey of Laboratories Across Pakistan. EJIFCC. 2025;36(2):143.
15. Jokic A, Rimac V, Vlasic Tanaskovic J, Podolar S, Honovic L, Lenicek Krleza J. The concurrence of the current postanalytical phase management with the national recommendations: a survey of the Working Group for Postanalytics of the Croatian Society of Medical Biochemistry and Laboratory Medicine. Biochemia medica. 2021;31(3):030704. DOI:10.11613/BM.2021.030704.
16. Tan RZ, Markus C, Loh TP. Relationship between biological variation and delta check rules performance. Clinical biochemistry. 2020;80:42-7. DOI:10.1016/j.clinbiochem.2020.03.017.
17. Ulas M, Yin J. B-108 Delta Check Performance Assessment by Using Real-WorldData at the Temple University Health System Clinical Chemistry.
18. Hong J, Cho EJ, Kim HK, Lee W, Chun S, Min WK. Application and optimization of reference change values for Delta Checks in clinical laboratory. Journal of clinical laboratory analysis. 2020;34(12):e23550. DOI:10.1002/jcla.23550.
19. Huf W, Mohns M, Garmatiuk T, Lister R, Buchta C, Ettl B, et al. Benchmarking diagnostic laboratory performance: Survey results for Germany, Austria, and Switzerland. Wien Klin Wochenschr. 2022;134(3-4):174-181. DOI:10.1007/s00508-021-01962-4.
20. Huf W, Mohns M, Bunning Z, Lister R, Garmatiuk T, Buchta C, et al. Benchmarking medical laboratory performance: survey validation and results for Europe, Middle East, and Africa. Clinical chemistry and laboratory medicine. 2022;60(6):830-841. DOI:10.1515/cclm-2021-1349.
21. Tan RZ, Markus C, Loh TP. Impact of delta check time intervals on error detection capability. Clinical chemistry and laboratory medicine. 2020;58(3):384-9. DOI:10.1515/cclm-2019-1004.
22. Liu L, Consagra W, Cai X, Mathias A, Worster A, Ma J, et al. Sex-Specific Absolute Delta Thresholds for High-SensitivityCardiac Troponin T. Clin Chem. 2022;68(3):441-449. DOI:10.1093/clinchem/hvab230.
23. Jeon BR. Blood urea nitrogen/creatinine change ratio as a delta check method for dialysis specimens. International journal of clinical and experimental pathology. 2020;13(7):1640-5.
24. Raza S, Pinkerton P, Hirsh J, Callum J, Selby R. The historical origins of modern international normalized ratio targets. J Thromb Haemost. 2024;22(8):2184-2194. DOI:10.1016/j.jtha.2024.05.013.
25. Zhou R, Liang YF, Cheng HL, Wang W, Huang DW, Wang Z, et al. A highly accurate delta check method using deep learning for detection of sample mix-up in the clinical laboratory. Clinical chemistry and laboratory medicine. 2022;60(12):1984-1992. DOI:10.1515/cclm-2021-1171.
26. Yang J, Wen S, McCudden CR, Tacker DH. Selection of Single-Analyte Delta Check Rules with Logistic Regression for Detection of Intravenous Fluid Contamination in a Clinical Chemistry Laboratory. The journal of applied laboratory medicine. 2024;9(5):1001-1013. DOI:10.1093/jalm/jfae066.
27. Choucair I, Lee ES, Vera MA, Drongmebaro C, El-Khoury JM, Durant TJS. Contamination of clinical blood samples with crystalloid solutions: An experimental approach to derive multianalyte delta checks. Clinica chimica acta; international journal of clinical chemistry. 2023;538:22-28. DOI:10.1016/j.cca.2022.10.011.
28. Colpaert K, Hoste EA, Steurbaut K, Benoit D, Van Hoecke S, De Turck F, et al. Impact of real-time electronic alerting of acute kidney injury on therapeutic intervention and progression of RIFLE class. Critical care medicine. 2012;40(4):1164-1170. DOI:10.1097/CCM.0b013e3182387a6b.
29. Flynn N, Dawnay A. A simple electronic alert for acute kidney injury. Annals of clinical biochemistry. 2015;52(Pt 2):206-212.DOI:10.1177/0004563214534832.
30. Kashani KB, Awdishu L, Bagshaw SM, Barreto EF, Claure-Del Granado R, Evans BJ, et al. Digital health and acute kidney injury: consensus report of the 27th Acute Disease Quality Initiative workgroup. Nat Rev Nephrol. 2023;19(12):807-818.10.1038/s41581-023-00744-7.
31. Elston DM. Participation bias, self-selection bias, and response bias. J Am Acad Dermatol. 2021. DOI:10.1016/j.jaad.2021.06.025.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Big Data and Diabetes: Seven-Year Data from Pakistan's Nation-Wide Laboratory Network

Lena Jafri[1*], Yousra Sarfaraz[1], Sibtain Ahmed[1], Zaib Un Nisa[1], Bilal Hashmi[1], Saad Bin Zafar Mahmood[1], Ayesha Sadiqa[1], Aysha Habib Khan[1], Imran Siddiqui[1], Hafsa Majid[1]

Medical College and Section of Chemical Pathology, Department of Pathology and Laboratory and 2nd Department of Medicine, Aga Khan University (AKU), Karachi

Article Info

[*]***Corresponding Author:***
Lena Jafri
Associate Professor, Section of Chemical Pathology Department of Pathology and Laboratory Medicine Aga Khan University, Karachi Pakistan
ORCID ID:0000-0001- 6807-5429
Contact number: +923222711537
E-mail: lena.jafri@aku.edu

Keywords

Diabetes Mellitus, Glucose, HbA1c, Low and middle income countries, Pakistan, Prediabetes, Prevalence

Abstract

Introduction: Diabetes has become a prominent worldwide public health issue, with its prevalence notably higher in lower middle-income countries. Pakistan ranks third globally in terms of number of individuals diagnosed with Type 2 Diabetes Mellitus and have highest prevalence of diabetes throughout the world. The objective of current study was to assess the frequency of diabetes in relation to demographics utilizing laboratory data of individuals residing in Pakistan.

Methods: Biochemical data of HbA1c, fasting and random plasma glucose, and oral glucose tolerance test were retrieved from Laboratory Information System analyzed in Pathology and Laboratory Medicine department from Jan 2016 to Dec 2022. Specimens were received from Aga Khan University laboratories' network, which is spread over the country. For interpretation of biochemical tests' values American Diabetes Association (ADA) criteria was applied.

Results: After removal of duplicates and missing data, the final sample included 4,260,171 unique individuals. Diabetes was present in 35.9% of the overall population, with the highest incidence (55.6%) recorded in males aged 35 to 59 years. Diabetes was significantly associated with advancing age and gender ($p<0.001$). Individuals who had
the fasting blood glucose ≥ 126 mg/dl were (n=372,050), HbA1c exceeding 6.5% were (n=762,940) the random blood glucose ≥ 200 mg/dl were (n=128,820), OGTT level ≥ 200 mg/dl were (n=487) and diabetes confirmed with more than one diagnostic test were (n=505,609).

Conclusion: The study, based on dataset of around 4.2 million individuals, revealed an overall diabetes frequency of 35.9%, with highest prevalence in middle-aged group (55.6%), and significant associations with both advancing age and gender.

Introduction

Globally, diabetes affected an estimated 463 million people in 2019, and the figure is expected to rise to around 700 million by 2045 [1]. It is a major public health concern worldwide, is more prevalent in low and middle -income countries (LMICs) such as Pakistan [2]. The International Diabetes Federation (IDF) reports that more than one in every four adults (26.7%) in Pakistan, a South Asian LMIC with a population of over 220 million, have Type 2 Diabetes Mellitus (T2DM) and has the highest national prevalence of diabetes throughout the world [3, 4]. Pakistan is expected to top the list in the near future, having surpassed China and India in terms of the number of people with T2DM globally [5].

The high prevalence of T2DM in LMICs is due to a complex interaction of socioeconomic, lifestyle, and environmental factors [6]. Rapid urbanization in LMICs has made a significant contribution to the rising T2DM rate. These changes include decreased physical activity, increased consumption of processed and high-calorie foods, and sedentary behaviors [7]. Traditional diets are frequently replaced by processed foods and high-fat, sugar-rich diets as countries transition economically. This nutritional shift increases the risk of obesity and, as a result, T2DM [8]. Genetic predisposition and family history are important factors in increasing diabetes risk in these populations [9, 10]. Furthermore, limited access to healthcare, lower education levels, and cultural norms that promote higher body weight aggravate the problem, making early detection and management of diabetes difficult [11]. Several developed countries, including the United States, Canada, Germany, Finland, Denmark, and Australia, have established diabetes registries; however, no such database exists in Pakistan. While global estimates exist, there is a lack of comprehensive national electronic registries to track real time diabetes prevalence in Pakistan. In the absence of comprehensive data, analyzing high-volume biochemical lab data could provide a preliminary understanding of diabetes burden in the country. The aim of this study involved utilizing the Aga Khan University Hospital's extensive laboratory network across Pakistan toexamine the period prevalence and gender distribution of diabetes and prediabetes across various age groups within the Pakistani populace.

Methodology

We curated and studied data of HemoglobinA1C (HbA1c), fasting and random blood glucose (FBG and RBG respectively) and Oral Glucose Tolerance Test (OGTT) in blood samples analyzed in the Section of Chemical Pathology, Department of Pathology and Laboratory Medicine. Data were extracted from the Laboratory Information System (LIS) of the Aga Khan University Hospital (AKUH). Presently, AKUH has a vast clinical laboratory network, serving the AKU health system, patients, and clinicians across Pakistan. The network includes laboratories on campuses and multiple collecting stations in over 90 districts with more than three hundred collection units, covering Pakistan's diverse population. The study was approved as an exemption by AKU's Ethical review committee (ERC number 2023-8559-24624). The data collection spanned from January 2016 to December 2022. Cases with missing or incomplete information and duplicate cases were removed. To maintain confidentiality, data encryption of patient identities was executed. Throughout the study, all HbA1c and glucose samples were tested using an ADVIA 1800, Siemens analyzer, New York, USA. The HbA1c assay has been certified by the National Glycohemoglobin Standardization Programme (NGSP), and its results can be directly linked or compared to those obtained using the Diabetes Control and Complications Trial (DCCT) reference method (for diagnostic purposes). Throughout the study, the laboratory was subject to proficiency testing by the College of American Pathologists. FBG, RBG and OGTT were performed by hexokinase method while HbA1c was analyzed using enzymatic assay method. In all biochemical analysis systems, performance was evaluated using three levels of quality control materials: low, medium, and high. The 75-g OGTT was performed according to international protocol with fasting and timed plasma glucose samples analyzed using validated analytical laboratory methods. Table 1 describes the cutoff values for HbA1c, FBG, RBG, and OGTT, as per the ADA guidelines [12,13]

Table 1: Prediabetes and Diabetes Diagnostic Criteria as per American Diabetes Association.

Diagnostic Test	Normal	Prediabetes	Diabetes
FBG, mg/dl	≤ 99	100 – 125	≥ 126
HbA1C, %	≤ 5.7	5.8 - 6.4	≥ 6.5
RBG, mg/dl	≤ 139		≥ 200
OGTT, mg/dl	≤ 139	140 – 199	≥ 200

Fasting is defined as no caloric intake for at least 8 hours. A1C tests should be performed in a laboratory using a method that is NGSP certified and standardised to the DCCT assay. RBG is not used to diagnose Prediabetes.

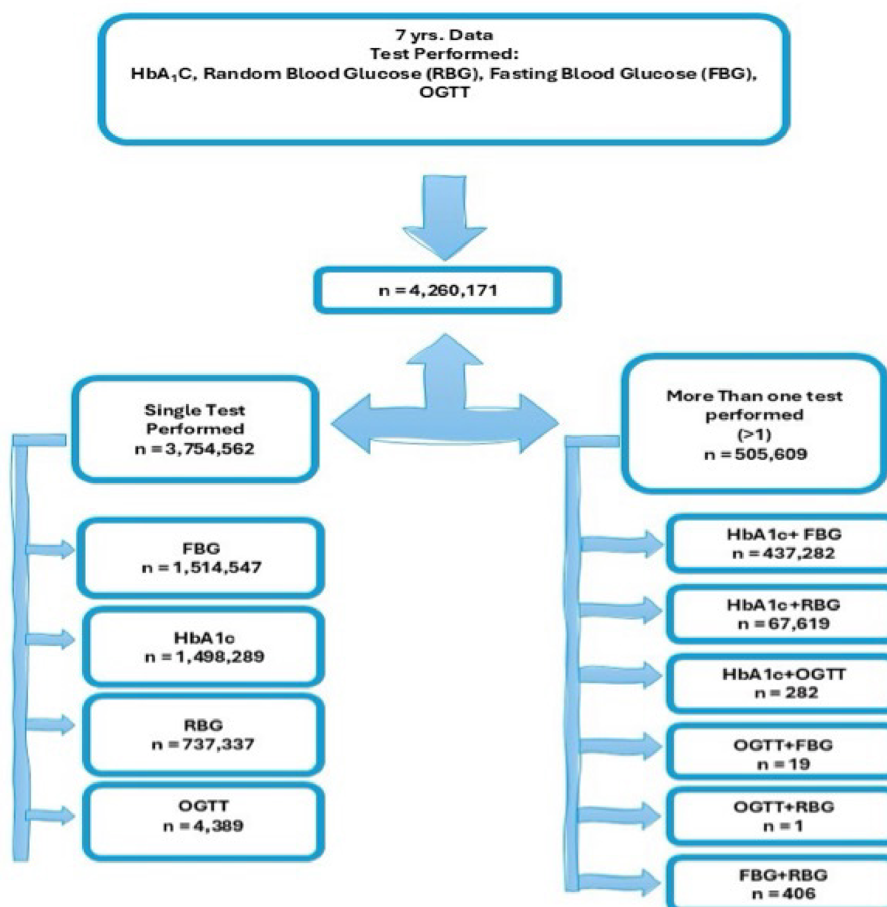
Statistical analysis

All data cleaning and analyses were performed on the software Stata version 17 (StataCorp LLC). Continuous variables values (FBG, HbA1C, RBG, OGTT) were reported using descriptive statistics, mean and standard deviation (SD). Using the ADA criteria above, age and gender variables were also grouped to construct a categorical variable. Categorical variables, including gender, were reported as frequencies and percentages. Age was categorized into four groups as follows: <18 years, 18 to 38 years, 39 to 59 years and >60 years. Chi-square test was used to calculate the statistical significance of difference in proportions amongst the categorical variables. Values were considered statistically significant when p value was<0.05. Additionally, Pakistan-specific spatial coordinates and divisions were found using the Global Administrative Areas Database (GADM) (www.gadm.org). The laboratory and hospital databases were used to construct spatial coordinates for collection sites whereblood specimens were given for analysis. Google Maps was used to gather correct spatial information in circumstances where details were absent, or verification was required. The study used Power BI software to create a provincial level choropleth map by aggregating frequencies and individual counts based on GADM's map and spatial subdivision of provinces.

Results

In this study, a total of 4,260,171 laboratory tests (HbA1c, FBG, RBG and OGTT) were studied. Figure 1 illustrates the sample categorization according to number of biochemical tests requested per individual. Out of the total biochemical tests analyzed, 88.1% were single tests and 11.9% were combination tests with the majority being single testing of HbA1c (n=1,498,289). The study included a total of 4,174,967 adult samples, accounting for approximately 98.0% of the total, with 85,204 children (<18 years) included, constituting about 2.0% of the samples.

Figure 1: Seven years data sample flow categorized according to diabetes diagnostic tests performed as part of panel and solo.



This flowchart illustrates the distribution of clinical samples over a seven-year period. Samples are categorized based on whether diabetes diagnostic tests were performed as individual (solo) assays or as part of a multitest panel. Classifications follow the ADA diagnostic criteria detailed in Table 1.

Out of the total (n=4,260,1 71), 35.9% (n=1,533, 523) had increased levels of biochemical markers encompassing cases through either a single positive marker or a combination of elevated HbA1c, RBG, FBG, and OGTT results indicating diabetes as described in Table 2. Out of the total tests conducted20.4% (n= 872,024) were identified as prediabetes either through either a single positive marker or a combination of elevated HbA1c, FBG, and OGTT results using ADA cutoffs. Out of total biochemical tests studied including the combination testing of the studied population had prediabetes. The majority

of diabetescases were identified through HbA1c testing (49.7%), followed by FBG testing (24.2%). The mean adult age among diabetes cases was 48 ±14.9 years and mean age of children were 10.2 ± 5.4 years. Age emerged as a crucial factor, with the highest frequency of diabetes and prediabetes observed in the 39-59 years age group as described in Table 3. Statistical analysis demonstrated a significant association between diabetes and both advancing age and gender, with *p*-values less than 0.001, highlighting the multifaceted nature of this disease's epidemiology. Frequency of diabetes and prediabetes was more in males (40.3% and 22% respectively) as compared to females (31.5% and 18.8% respectively); *p* value <0.001. Table 4 denotes the Multinomial regression of diabetes and prediabetes showing significant correlation between age and gender for this cohort. For diabetes, the factor Age (RRR = 1.0562): One unit of age increases the relative risk of diabetes by 5.62%, compared to the normal group (p < 0.001) while in factor gender for Male (RRR = 1.4445) indicating, Males have a 44.45% higher risk of diabetes compared to females, controlling for age (p < 0.001). Similarly for Prediabetic category, for Age (RRR = 1.0439): One unit increase in age is associated with a 4.39% increase in the relative risk of being in the prediabetes group compared to the normal group (p < 0.001). As for the gender, Males (RRR = 1.1998): have a 19.98% higher relative risk of prediabetes than females.

Table 2: Total number of diabetes tests (HBA1C, OGTT, RBG, FBG) ordered categorized based on number of normal, prediabetes and diabetes ratio.

Biochemical Test	Overall n (%)	Normal n (%)	Prediabetes n (%)	Diabetes n (%)
Overall	4,260,171	1,854,624 (43.5)	872,024 (20.4)	1,533,523 (35.9)
FBG	1,514,547 (35.5)	767,648 (41.3)	374,849 (42.9)	372,050 (24.2)
HbA1C	1,498,289 (35.1)	374,052 (20.1)	361,297 (41.4)	762,940 (49.7)
RBG	737,337 (17.3)	608,517 (32.8)		128,820 (8.4)
OGTT	4,389 (0.1)	2,690 (0.14)	1,212 (0.13)	487 (0.03)
More than 1 test	505,609 (11.8)	101,717 (5.4)	134,666 (15.4)	269,226 (17.5)

For single-test evaluations, FBG, OGTT, HbA1c were treated as standard diagnostic markers. Whereas n is the number of test performed or total number of normal, diabetic and prediabetic results and % is total number of test (HBA1C/ OGTT/ RBG/ FBG) / overall test performed and total number of normal, diabetic and prediabetic results / total number of HBA1C/ OGTT/ RBG/ FBG performed respectively.

Table 3: Distribution of Prediabetes and Diabetes by Age and Gender in Clinical Laboratory Data Over Seven Years.

Age Category	Female	Female	Female	Male	Male	Male	p- value
Age Category	Total n(%)	Prediabetes n(%)	Diabetes n(%)	Total n(%)	Prediabetes n(%)	Diabetes n(%)	p- value
Overall	2,091,598 (49.0)	394,416 (18.8)	659,544 (31.5)	2,168,573 (50.9)	477,608 (22)	873,979 (40.3)	<.001
Adolescent (<18)	35,762 (1.7)	2,277 (0.57)	5,633 (0.85)	37,909 (1.7)	2,758 (0.57)	6,018 (0.6)	<0.001
Early Adulthood (18-38)	676,243 (32.3)	60,109 (15.2)	69,700 (10.5)	448,742 (20.7)	77,215 (16.1)	105,555 (12.0)	<0.001
Middle Age (39-59)	877,922 (41.9)	196,147 (49.7)	354,054 (53.6)	1,097,456 (50.6)	252,271 (52.8)	486,543 (55.6)	<0.001
Old Age (≥ 60)	501,671 (23.9)	122,570 (31.0)	230,157 (34.9)	584,466 (26.9)	145,364 (30.4)	275,863 (31.5)	<0.001

*Chi square test significant association **(p<0.001)** is used, in female, diabetes and prediabetes are significantly associated with age categories. As well as male also show significant difference between age categories with prediabetes and diabetes*

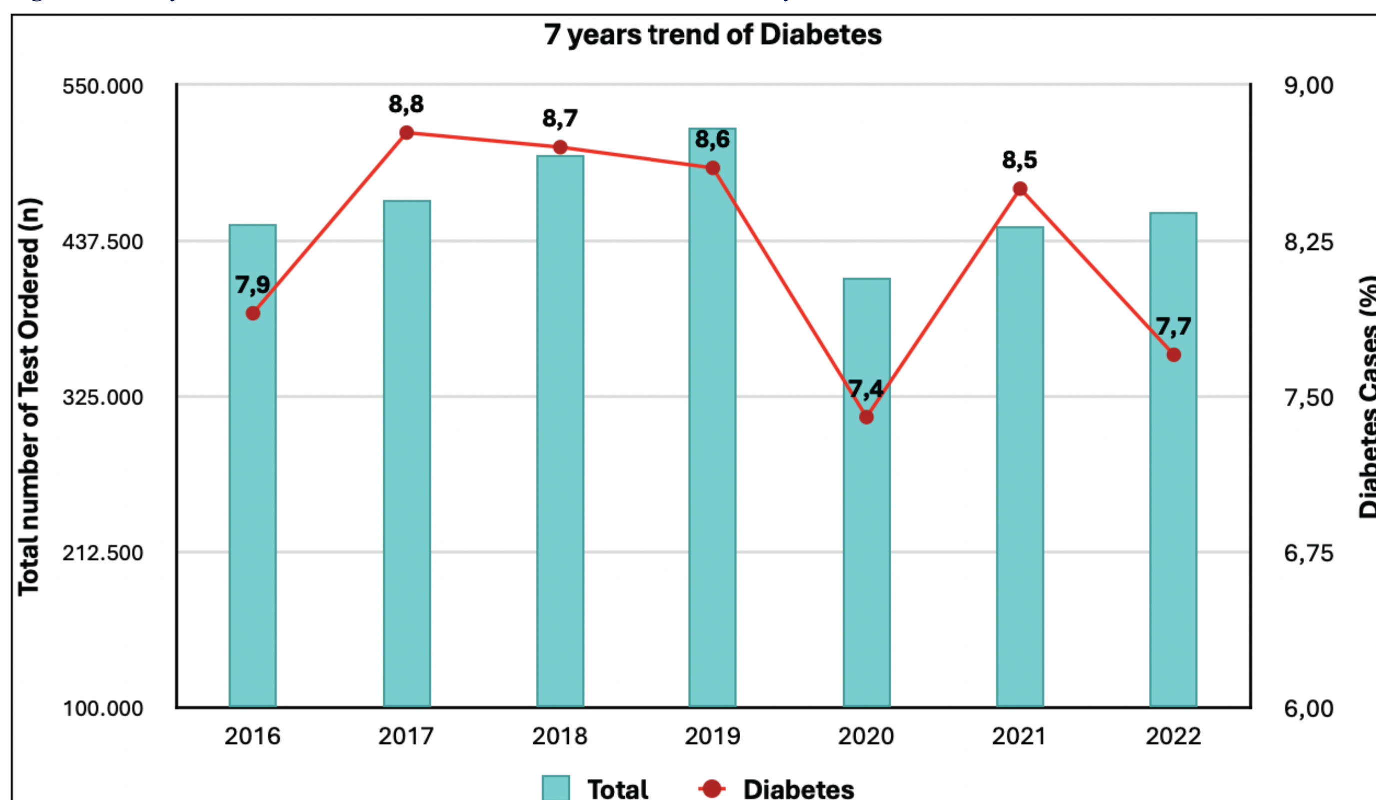
Table 4: Multinomial Regression of Diabetes and Prediabetes Data, Showing Significance between Age and Gender.

Outcome Category	Variable	RRR	Std. Error	z-value	p-value	95% CI Lower	95% CI Upper
Prediabetes	Age	1.044	0.000	251.100	<0.001	1.044	1.044
	Male (ref: Female)	1.200	0.006	37.980	<0.001	1.189	1.211
	Constant	0.119	0.001	−247.05	<0.001	0.117	0.121
Diabetes	Age	1.056	0.000	356.020	<0.001	1.056	1.057
	Male (ref: Female)	1.445	0.006	87.080	<0.001	1.433	1.457
	Constant	0.123	0.001	−272.59	<0.001	0.121	0.125

Age significantly increases the risk of both prediabetes and diabetes. Males are significantly higher risk than females for both conditions. All predictors are statistically significant at p<0.001.

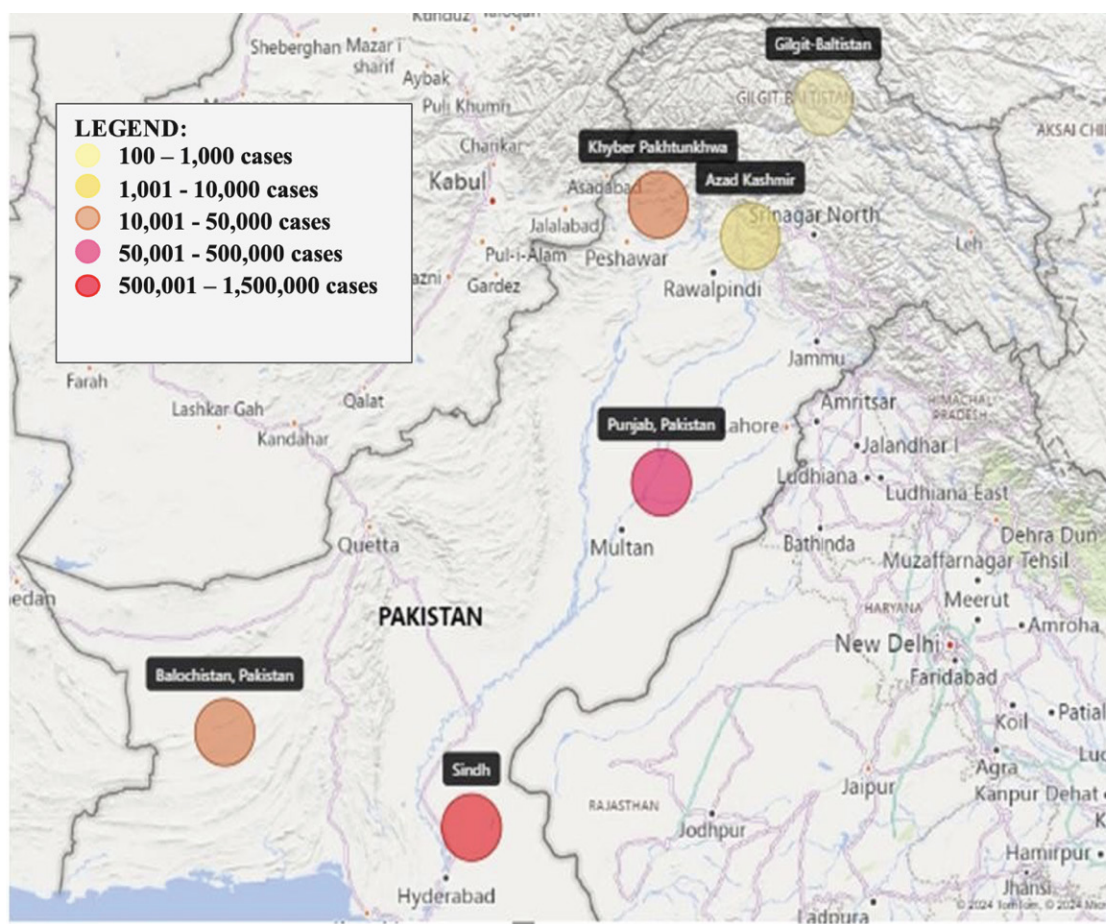
Figure 2 denotes the trend of diabetes Figure 2 denotes the trend of diabetes within country; the graph shows the constant rise of the diabetes case till 2019 maintaining the fluctuation and hike at least from recent years. The choropleth map (Figure 3) shows the cases with diabetes from different regions of Pakistan indicating 214 cases of diabetes were reported from Gilgit-Baltistan, 4,918 cases from Azad Jammu Kashmir (AJK), 46,905 cases from Balochistan, 45,952 cases from Khyber Pakhtunkhwa (KPK), 351,221 cases from Punjab and 1084,313 cases from Sindh. There were additional samples with missing or inadequate location data that could not be traced.

Figure 2: Yearly Pattern of Diabetes Cases Identified in Clinical Laboratory Data Over Seven Years.



Cases in Figure 2 were identified from clinical samples across Pakistan and analyzed at a central laboratory. Classification was based on American Diabetes Association (ADA) diagnostic criteria, as detailed in Table 1.

Figure 3: Choropleth Map Showing Distribution of Diabetes Cases by Province in Clinical Laboratory Data Over Seven Years.



This map illustrates the prevalence of diabetes across provinces based on clinical laboratory records. Cases were identified using ADA diagnostic criteria: Color intensity represents the concentration of cases, with darker shading indicating a higher volume of identified individuals. Data were processed in a central clinical laboratory from samples collected nationwide.

Discussion

In this study, conducted on a sample set of approximately 4.2 million laboratory specimens, the period prevalence of diabetes emerged as a significant health concern, with 35.9% of the cases affected. Our data depicted that males are at higher diabetes rate (40.3%) than female while in terms of age, a higher rate of diabetes was observed in middle-aged men (55.6%) as compared to middle aged females (53.6%). In prediabetics, similar ratio was observed with males having slightly higher ratio than females with the highest prediabetic ratio seen in middle-aged male group. This data analysis found that gender and age were significant predictors of hyperglycemia and prediabetes, which can be considered a risk factor for diabetes. The COVID-19 pandemic most likely attributed to a temporary drop in the trend line of diabetes prevalence in 2020 when routine screening rates declined followed by an unstable recovery.

The IDF Diabetes Atlas 2022, Pakistan ranks number one, with highest prevalence of diabetes 31%, with projection of prevalence over 34% keeping it at the top position in 2045 (3). The Pakistan National Diabetes Survey showed that in urban areas 9.83% of women and 12.14% of men had diabetes. Furthermore, impaired glucose tolerance was present in 9.54% of females and 4.54% of males [14]. Studies indicated that the prevalence of glucoseintolerance is high in the Pakistani population. Local data assessed from current studies also showed that Sindh and Punjab have the highest reported diabetes cases and similarly have higher risk prevalence of diabetes [15]. According to studies prevalence of diabetes was 26.3%, with 19.2% known and 7.1% newly diagnosed. Urban areas have 28.3% prevalence, while pre-diabetes is 14.4%, the reported mean age (greater than or equal to 43 years) was similar coinciding with our studies [16]. Key factors for higher period prevalence of diabetes in Sindh and Punjab include urbanization causing sedentary lifestyles and unhealthy diets, high obesity rates, and the shift to urbanization. Cities like Karachi, Hyderabad, Lahore, Islamabad, Multan and Faisalabad have better access to healthcare facilities and clinical laboratories. A larger middle and upper class in these provinces can afford health check-ups, and many health awareness programs are active in the region.

The National Diabetes Survey shows an increase in type 2 diabetes and Impaired Glucose Tolerance (IGT) prevalence with age in both sexes, consistent with other South Asian studies [17] it further indicates that at national level while the prevalence of impaired glucose tolerance (IGT) is higher in women than in men, the prevalence of diabetes is higher in men. Diabetes in Pakistan women presents a complex, multidimensional scenario.

Sociocultural barriers, limited healthcare access, and the metabolic stress of multiparity risk higher morbidity and mortality rates. Gestational
Diabetes Mellitus (GDM) can serve as a blessing in disguise for early diagnosis and intervention. It shall sensitize multi-stakeholders for prevention of T2DM focusing on a healthy lifestyle at comparatively younger age group, and in turn also addressing interventions for child hood obesity, GDM shall be integrated into maternal and child health (MCH) care programs [18, 19].
Subsequently, the laboratory data showed that even children (<18) were also diagnosed with diabetes and at-risk status (pre diabetes) was also observed, though the data availability was limited but it was still indicative of the impelled effects of diabetes. A study conducted in 2010 estimated that though the actual number of diabetic children within the country is unknown still it was approximated that around 61,196 children are affected with an increment in this figure in every passing year [20]. Diabetes risk in children increases with high BMI, blood pressure, and larger waists, influenced by factors like genetic predisposition, urbanization, lifestyle changes, obesity, malnutrition, depression, and socio-cultural issues [21,22]. The country's choropleth map for diabetes highlights significant regional disparities in the reported cases of diabetes. This geographic distribution emphasizes the importance of focused public health measures for diabetes, including healthy diet, regular exercise, diabetes education, screenings, and healthcare access in the most affected areas, primarily in Sindh and Punjab. These findings raise important questions about the underlying factors contributing to the regional disparities in diabetes prevalence. More large-scale longitudinal research is needed to answer the question of whether environmental, socioeconomic, or lifestyle factors are driving these developments in various provinces of Pakistan.
The study conducted has its own limitations as it is a cross sectional study, anthropometric measurements, clinical history, lifestyle factors and dietary intake were not available and no follow up was carried out on these individuals. The fact that this article includes a significant number of patients who represent both urban and rural populations makes it possible to generalize the findings throughout Pakistan, which is one of its strengths. The choropleth map also clearly illustrates the disease burden, emphasizing the necessity of focused interventions. The absence of a specific database or registry is still a drawback, though.

Conclusion and Call to Action

According to the data, age is positively associated with both prediabetes and diabetes. Males are at significantly higher risk for both prediabetes and diabetes compared to females. Middle-aged men reported more positive cases, which means they are at higher prevalence of risk. Nonetheless, the percentage of females is not far behind, and a significant number of cases have also been reported among women. Sindh and Punjab showed the higher risk of diabetes while may be due to data unavailability from their provinces the Pakistan population tendency to become diabetic cannot be ruled out. As age increases, non-communicable diseases increase, necessitating efficient and diligent treatment.
Human resources and evidence-based solutions are crucial, prompting multi-stakeholder action.
The current study highlights the need for establishing a national multi-center diabetes database using identical, consistent definitions and data collection methods for prevalence and incidence of diabetes across different times, locations, and standardized frequency measures. To facilitate robust cross-study comparisons and ensure the validity of longitudinal analyses, it is imperative to establish a standardized diabetes reporting framework for prevalence and incidence. A key component of the creation of a comprehensive digital database for both pre-diabetics and diabetics, can also be promoting active, real-time management and facilitating personalized lifestyle modifications and care.

Authors' contributions

YS substantially contributed to the conception, design of the work, data acquisition, analysis and interpretation, and drafted and revised the paper. LJ designed and conceptualized the work, acquired the data, and contributed to drafting and revising the paper. SA, ZN, AS substantively analyzed and made significant contributions to the revision of the work. BH, SBZN revised the data analysis and manuscript. AS primarily works on data decoding and statistical analysis. AH, IS and HM contributed to manuscript revision. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Disclosure of Ethical Statements

Approval of the research protocol: An exemption has been acquired from the institutional ethical review committee (Exemption number 2023-8559-24624).
Informed Consent: N/A
Approval date of Registry and the Registration No. of the study/trial: N/A Animal Studies: N/A

Disclosure

The authors declare that they have no competing interests, and No funding was received for conducting this study.

References

1. Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Colagiuri S, Guariguata L, Motala AA, Ogurtsova K, Shaw JE. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. Diabetes research and clinical practice. 2019;157:107843. DOI: 10.1016/j.diabres.2019.107843.
2. Bhutta ZA, Haq ZU, Basit A. Diabetes in Pakistan: addressing the crisis. The Lancet Diabetes & Endocrinology. 2022;10(5):309-310. DOI: 10.1016/S2213-8587(22)00102-4.
3. Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, Stein C, Basit A, Chan JC, Mbanya JC, Pavkov ME. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes research and clinical practice. 2022;183:109119. DOI: 10.1016/j.diabres.2021.109119.

4. Diabetes responsible for 400,000 deaths in Pakistan in 2021: Report (app.com.pk), available from https://www.app.com.pk/national/dia betes-responsible-for-400000-deaths-in-pakistan-in-2021-report/. (assessed: 05/07/2024).
5. Hossain MJ, Al□Mamun M, Islam MR. Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused. Health Science Reports. 2024;7(3):e2004. DOI: 10.1002/hsr2.2004.
6. Richards SE, Wijeweera C, Wijeweera A. Lifestyle and socioeconomic determinants of diabetes: Evidence from country-level data. PLoS One. 2022;17(7):e0270476. DOI: 10.1371/journal.pone.0270476.
7. Zahid N, Claussen B, Hussain A. Diabetes and impaired glucose tolerance in a rural area in Pakistan and associated risk factors. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2008;2(2):125-130. DOI: 10.1016/j.dsx.2008.02.007.
8. Popkin BM, Ng SW. The nutrition transition to a stage of high obesity and noncommunicable disease prevalence dominated by ultra□processed foods is not inevitable. Obesity Reviews. 2022;23(1):e13366. DOI: 10.1111/obr.13366.
9. InterAct Consortium robert. scott@ mrc-epid. cam. ac. uk. The link between family history and risk of type 2 diabetes is not explained by anthropometric, lifestyle or genetic risk factors: the EPIC-InterAct study. Diabetologia. 2013;56:60-69. DOI: 10.1007/s00125-012-2715-x.
10. Mathias RA, Deepa M, Deepa R, Wilson AF, Mohan V. Heritability of quantitative traits associated with type 2 diabetes mellitus in large multiplex families from South India. Metabolism. 2009;58(10):1439-1445. DOI: 10.1016/j.metabol.2009.04.041.
11. Adu MD, Malabu UH, Malau-Aduli AE, Malau-Aduli BS. Enablers and barriers to effective diabetes self-management: A multi-national investigation. PloS one. 2019;14(6):e0217771. DOI: 10.1371/journal.pone.0217771.
12. Song WM, Liu Y, Zhang QY, Liu SQ, Xu TT, Li SJ, An QQ, Liu JY, Tao NN, Liu Y, Yu CB. Ambient air pollutants, diabetes and risk of newly diagnosed drug-resistant tuberculosis. Ecotoxicology and environmental safety. 2021;219:112352. DOI: 10.1016/j.ecoenv.2021.112352.
13. American Diabetes Association. 2. Classification and diagnosis of diabetes. Diabetes care. 2016;39(Supplement_1):S13-22. DOI: 2337/dc16-S005.
14. Shera AS, Basit A, Fawwad A, Hakeem R, Ahmedani MY, Hydrie MZ, Khwaja IA. Pakistan National Diabetes Survey: prevalence of glucose intolerance and associated factors in the Punjab Province of Pakistan. Primary care diabetes. 2010;4(2):79-83. DOI: 10.1016/j.pcd.2010.01.003.
15. Akhtar S, Nasir JA, Abbas T, Sarwar A. Diabetes in Pakistan: A systematic review and meta-analysis. Pakistan journal of medical sciences. 2019;35(4):1173. DOI: 10.12669/pjms.35.4.194.
16. Basit A, Fawwad A, Qureshi H, Shera AS. Prevalence of diabetes, pre-diabetes and associated risk factors: second National Diabetes Survey of Pakistan (NDSP), 2016–2017. BMJ open. 2018;8(8):e020961. DOI: 10.1136/bmjopen-2017-020961corr1.
17. Shah A, Kanaya AM. Diabetes and associated complications in the South Asian population. Current cardiology reports. 2014;16:1-6. DOI: 10.1007/s11886-014-0476-5.
18. Shera AS, Jawad F, Maqsood A. Prevalence of diabetes in Pakistan. Diabetes research and clinical practice. 2007;76(2):219-222. DOI: 10.1016/j.diabres.2006.08.011.
19. Dunstan DW, Zimmet PZ, Welborn TA, De Courten MP, Cameron AJ, Sicree RA, Dwyer T, Colagiuri S, Jolley D, Knuiman M, Atkins R. The rising prevalence of diabetes and impaired glucose tolerance: the Australian Diabetes, Obesity and Lifestyle Study. Diabetes care. 2002;25(5):829-834. DOI: 10.2337/diacare.25.5.829.
20. Qayyum AA, Lone SW, Ibrahim MN, Atta I, Raza J. Effects of diabetes self-management education on glycaemic control in children with insulin-dependent diabetes mellitus. J Coll Physicians Surg Pak. 2010;20(12):802-805. PMID: 21205545.
21. Rashid AS, Jabbar A, Michels RP, DeVries JH. Metabolic risk factors, insulin-resistance and lifestyle in children of type 2 diabetes patients in Karachi, Pakistan. Diabetes research and clinical practice. 2008;80(3):399-404. DOI: 10.1016/j.diabres.2008.01.013
22. Basit A, Riaz M, Fawwad A. Improving diabetes care in developing countries: The example of Pakistan. Diabetes research and clinical practice. 2015;107(2):224-232. DOI: 10.1016/j.diabres.2014.10.013.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Bisalbuminemia: A rare finding identified via serum protein agarose gel electrophoresis and capillary electrophoresis techniques

Qian Ren[1,2], Tong Liu[3], Jing Li[1,2], Yulong Fan[1,2], Shoulei Wang[1, 2], Lele Wang[1,2], Yansheng Wang[1,2], Zhaojing Liu[1,2], Yonghui Xia[1,2], Guoqing Zhu[1,2], Yansong Ren[1,2*]

State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China

[2]Tianjin Institutes of Health Science, Tianjin, China

Article Info

Corresponding Author:

Yansong Ren

State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 288, Nanjing Road, Heping District, Tianjin, 300020, China.

Phone.: +86 022-23909181

E-mail: renyansong@ihcams.ac.cn

Keywords

Bisalbuminemia, Agarose gel electrophoresis, Capillary serum protein electrophoresis

Abstract

Background: Bisalbuminemia is a rare abnormality characterized by detection of two distinct albumin bands or double peaks through serum protein agarose gel electrophoresis or capillary electrophoresis, suggesting the coexistence of normal and modified albumin in the same patient. However, it is challenging to visualize the double bands due to the limited resolution of agarose gel electrophoresis. Besides, capillary electrophoresis cannot directly visualize serum proteins.

Methods: Serum samples were obtained from the Hematology Hospital of the Chinese Academy of Medical Sciences and analyzed using agarose gel electrophoresis and capillary electrophoresis. Agarose gel electrophoresis and capillary electrophoresis were performed using a SPIFE4000 electrophoresis apparatus from Helena or Capillarys3 Tera system.

Results: Three samples of bisalbumin were detected in the albumin region in patients diagnosed with multiple myeloma via agarose gel electrophoresis and capillary electrophoresis techniques.

Conclusions: These findings suggest that capillary electrophoresis technology combined with traditional agarose gel electrophoresis methods can significantly improve the detection rate of bisalbuminemia. Serum dilution or depolymerization may be utilized to enhance detection when bisalbumin bands are not readily identifiable in agarose gel electrophoresis, thereby ensuring the accuracy of bisalbuminemia diagnosis.

Introduction

Human serum albumin is synthesized by hepatocytes in the liver. Serum albumin consists of a single polypeptide chain comprising 585 amino acids and has a molecular weight of 66.5 kDa. It has multiple physiological functions, including sustaining plasma colloid osmotic pressure, inhibiting the formation of oxidants, catalyzing biochemical reactions, and transporting endogenous and exogenous substances (thyroid hormones, steroids, and fatty acids) [1-2].

Clinical chemistry analyzers are widely used to quantify serum albumin in clinical practice. Besides, protein components of albumin are identified through serum protein electrophoresis. Agarose gel electrophoresis is one of the simplest and widely utilized methods for the separation of serum proteins in clinics due to its distinct banding patterns and high reproducibility. Using this technique, researchers can differentiate serum proteins on the basis of variations in their isoelectric points and molecular weights. The proteins can be grouped into several categories, including albumin, alpha-1 globulins, alpha-2 globulins, beta globulins, and gamma globulins. However, capillary electrophoresis provides rapid detection and enhanced resolution compared with traditional agarose gel electrophoresis due to the advancement of electrophoresis technology [3]. Capillary electrophoresis separates proteins based on their charges and sizes, resulting in varying migration velocities within an electric field [4]. The detector captures the migration speeds of various protein components, producing electropherograms with different electrophoretic peaks. The protein components, including albumin, alpha-1 globulins, alpha-2 globulins, beta-1 globulins, beta-2 globulins, and gamma globulins, can be identified by analyzing these peaks [5]. The detector can be used for serum protein electrophoresis and immunotyping detection in myeloma. Notably, the methods are critical for the screening and diagnosis of multiple myeloma.

Normal human serum albumin appears as a single band or a prominent electrophoretic peak during serum protein detection. However, two distinct bands may be detected in the albumin region in rare cases during serum protein agarose gel electrophoresis, a phenomenon known as bisalbuminemia [6]. This condition was first described in diabetic patients by Scheurlen in 1955, who exhibited two distinct bands with different electrophoretic mobilities on agarose gel electrophoresis [1]. Bisalbuminemia occurs in about 1 in 1000-1 in 3000 people. However, the incidence of bisalbuminemia among North American Indians is about 1 in 100 individuals [9]. Bisalbuminemia can either be hereditary or acquired [7,8]. Hereditary bisalbuminemia is a permanent condition inherited in an autosomal codominant manner. In contrast, acquired bisalbuminemia is associated with certain diseases, such as diabetes or multiple myeloma, or in patients treated with penicillin and other drugs[10]. In this study, double electrophoretic bands and peaks were detected in the albumin region of three patients diagnosed with multiple myeloma using serum protein agarose gel electrophoresis and capillary electrophoresis techniques. Compared with individual detection methods, the integration of the above techniques facilitates the rapid visualization of double protein bands and enhances the sensitivity of detecting bisalbumin and electrophoretic peaks. This comprehensive approach promotes accurate and timely identification of bisalbuminemia.

Materials and Methods

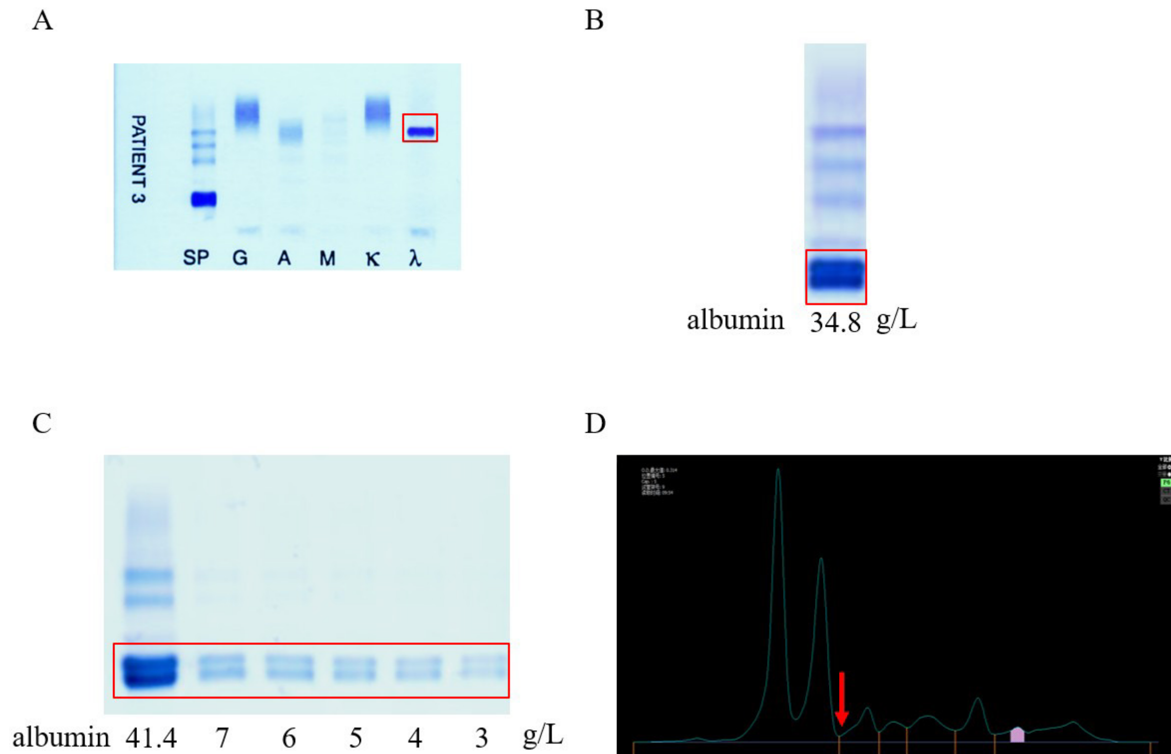
Three serum samples were obtained from the Hematology Hospital of the Chinese Academy of Medical Sciences and analyzed using agarose gel electrophoresis and capillary electrophoresis. Acquired causes of bisalbuminemia are not ruled out in the above cases. Agarose gel electrophoresis was conducted on the SPIFE4000 electrophoresis apparatus purchased from Helena to examine and results were quantified using an optical densitometer. Capillary electrophoresis was carried out using a fully automated Capillarys3 Tera system. The results were analyzed using PHORESIS software. Notably, serum samples were either diluted with normal saline or treated with ß-mercaptoethanol.

Results

Agarose gel electrophoresis and capillary electrophoresis can identify bisalbuminemia

An 82-year-old male patient with anemia admitted to our hospital was diagnosed with multiple myeloma following a bone marrow examination. Immunofixation electrophoresis identified a monoclonal lambda light chain component in the ß region (Figure 1A). Agarose gel electrophoresis detected double bands in the albumin region before treatment (Figure 1B). The patient underwent four cycles of the VRD (bortezomib, lenalidomide, and dexamethasone) regimen, resulting in a complete response. Subsequently, the patient received regular infusions of bisphosphonates to mitigate bone destruction and continued treatment with lenalidomide and ixazomib. However, treatment was adjusted to daratumumab upon re-evaluation due to disease progression. Agarose gel protein electrophoresis detected double bands in the albumin region during the treatment process. Notably, the bisalbumin bands were more distinct at a protein concentration of 7 g/L following concentration gradient dilution of the serum sample (Figure 1C). Similarly, capillary electrophoresis detected bisalbumin peaks (Figure 1D). These findings suggest that agarose gel electrophoresis and capillary electrophoresis can identify bisalbuminemia.

Figure 1: Agarose gel electrophoresis and capillary electrophoresis techniques detecting bisalbumin.

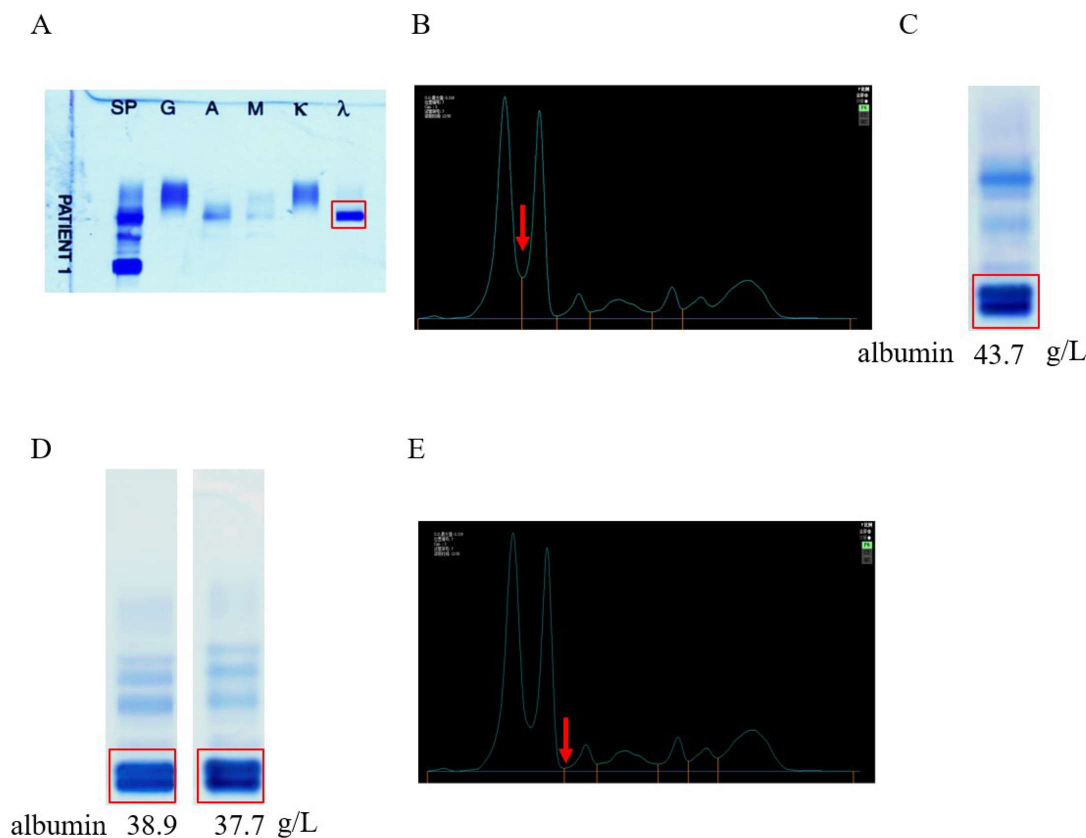


A. IFE showing a monoclonal light chain λ component in the β region. B. Agarose gel electrophoresis showing double bands in the albumin fraction before treatment. C. Agarose gel electrophoresis showing double bands in the albumin fraction after treatment. D. Capillary electrophoresis showing double peaks after treatment.

Capillary electrophoresis alone can easily result in misjudgment, which can be corrected via agarose gel electrophoresis

A 44-year-old male patient with elevated serum uric acid and creatinine levels presented to the nephrology department due to an unusual oral odor. Serum immunofixation electrophoresis revealed the presence of precipitated bands in the lambda lane. Further evaluation through a whole-body CT scan confirmed the presence of myeloma lesions. Besides, immunofixation electrophoresis demonstrated a monoclonal light chain λ component in the γ region, thereby confirming the diagnosis of multiple myeloma (Figure 2A). The patient was treated with Daratumumab combined with pomalidomide, bortezomib, and dexamethasone (VPD) regimen. The patients received continuous treatment with regular follow-ups. Notably, the firstelectrophoretic peak and second peak were initially misidentified as albumin and alpha-1 globulins, respectively, during the diagnostic and treatment process via capillary electrophoresis (Figure 2B). However, we retrospectively reviewed the agarose gel electrophoresis data from pre- and post-treatment and identified distinct double bands in the albumin region (Figure 2C-2D). The results of capillary electrophoresis were corrected, revealing the presence of bisalbumin electrophoretic peaks (Figure 2E). Capillary electrophoresis can only display peak patterns. However, expertise is required to recognize the difference, making it challenging to identify bisalbumin peaks. In contrast, agarose gel electrophoresis can accurately display two bands in the albumin region, thereby facilitating a more precise bisalbuminemia diagnosis.

Figure 2: Capillary electrophoresis may lead to erroneous interpretations.

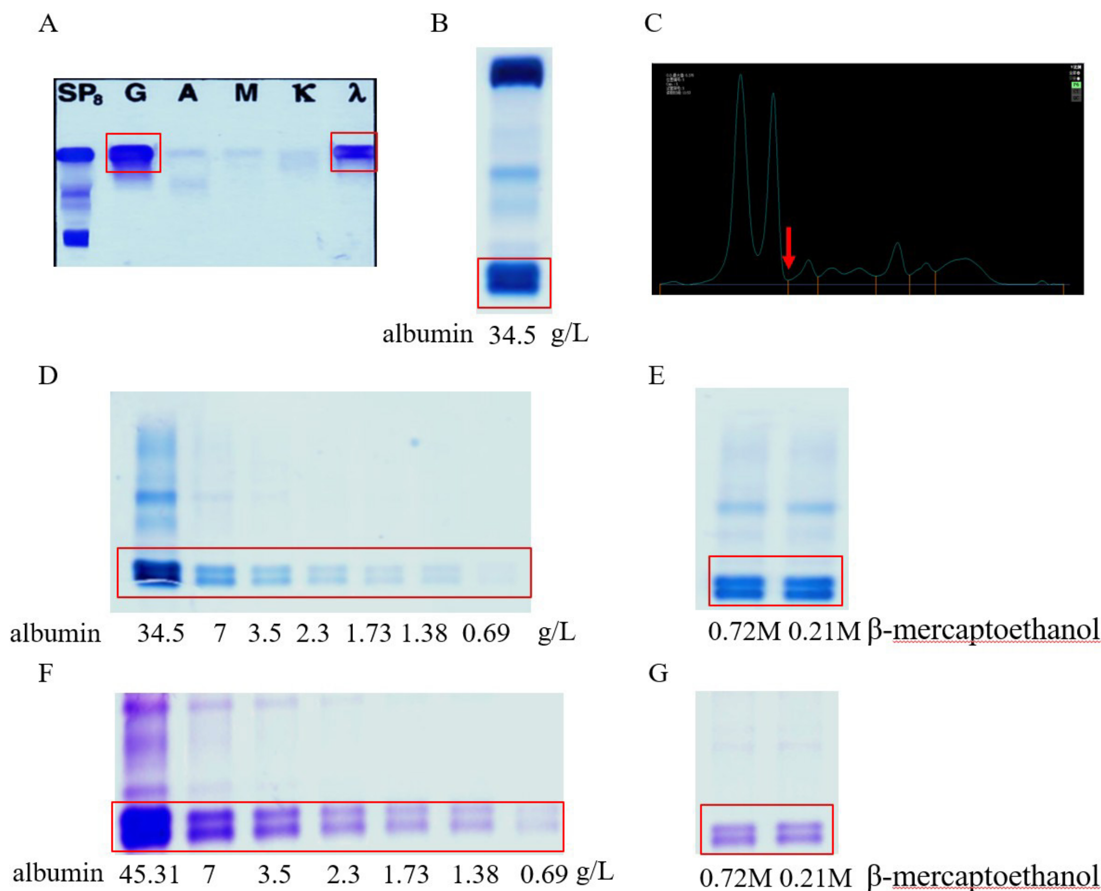


A. IFE showing a monoclonal light chain λ component in the γ region. B. The erroneous result of capillary electrophoresis where the initial electrophoretic peak was characterized as albumin, whereas the subsequent peak was classified as alpha-1 globulins after treatment. C. Agarose gel electrophoresis showing double bands within the albumin fraction before treatment. D. Agarose gel electrophoresis showing double bands within the albumin fraction after treatment. E. Capillary electrophoresis demonstrating the presence of double peaks in the albumin region after treatment.

The integration of agarose gel electrophoresis with capillary electrophoresis may facilitate a precise and expedited diagnosis of bisalbuminemia

A 49-year-old male patient with anemia, presented with foamy urine at a local hospital, and was diagnosed with multiple myeloma. The patient was admitted to our hospital for further evaluation and treatment. A thorough diagnostic evaluation, which included serum immunofixation electrophoresis, revealed the presence of a monoclonal IgG lambda protein in the γ region (Figure 3A). Key findings included an M-spike of 144.44%, an M-protein concentration of 48.806 g/L, and significantly elevated free lambda light chains at 4269.18 mg/L. These results indicate the presence of a monoclonal gammopathy, consistent with an abnormal proliferation of plasma cells. Further diagnosis was achieved through bone marrow aspiration and biopsy, flow cytometry, and screening for gene mutations associated with multiple myeloma. The patient underwent four cycles of chemotherapy utilizing a regimen with VPD. Agarose gel electrophoresis results revealed a concentrated band in the albumin region before treatment (Figure 3B). Additionally, capillary electrophoresis revealed the presence of double peaks in the albumin region (Figure 3C). The methodology was refined based on these observations. Quantitative analysis showed that the concentration of serum albumin was 34.5 g/L, with two distinct bands detected in the albumin region following concentration gradient dilution. The albumin bands were optimally separated at a serum albumin concentration of 7 g/L. However, the visibility of the bisalbumin bands diminished as the dilution ratio increased before treatment (Figure 3D). Protein depolymerization using 0.72 M and 0.21 M ß-mercaptoethanol corroborated this phenomenon. Notably, the depolymerization effect of ß-mercaptoethanol was not significantly different between the two concentrations (Figure 3E). We further employed agarose gel electrophoresis to analyze the serum samples obtained from patient after treatment and observed the same bisalbumin bands (Figure 3F-3G). These findings suggest that bisalbumin bands can be effectively observed through agarose gel electrophoresis at 7 g/L albumin concentration or when depolymerization is conducted using ß-mercaptoethanol at 0.72 M or 0.21 M. Therefore, the integration of agarose gel electrophoresis and capillary electrophoresis can provide a rapid and accurate diagnostic approach for bisalbuminemia.

Figure 3: The integration of agarose gel electrophoresis and capillary electrophoresis facilitates precise bisalbumin diagnosis.



A. IFE showing a monoclonal IgG lambda protein in the γ region.

B. Agarose gel electrophoresis demonstrating a concentrated band within the albumin fraction before treatment.

C. Capillary electrophoresis showing double peaks after treatment.

D. The quantified concentrations of albumin (34.5 g/L, 7 g/L, 3.5 g/L, 2.3 g/L, 1.73 g/L, 1.38 g/L and 0.69 g/L) before treatment.

E. Serum protein agarose gel electrophoresis performed after depolymerization using 0.72M and 0.21M β-mercaptoethanol before treatment.

F. The quantified concentrations of albumin (45.31 g/L, 7 g/L, 3.5 g/L, 2.3 g/L, 1.73 g/L, 1.38 g/L and 0.69 g/L) after treatment.

G. Serum protein agarose gel electrophoresis performed after depolymerization using 0.72M and 0.21M β-mercaptoethanol after treatment.

Discussion

Bisalbuminemia is characterized by the presence of two distinct albumin detected through serum protein electrophoresis, suggesting a variant albumin in the serum. This condition is associated with the occurrence of diverse diseases [11-14]. Compared with normal albumin, these variants exhibit functional alterations, such as decreased binding affinity to molecules (testosterone and bilirubin), a shorter half-life, and suppressed binding to warfarin. Notably, other variants demonstrate increased binding affinity to progesterone, triiodothyronine (T3), thyroxine (T4), and fatty acids [2,15]. However, acquiredbisalbuminemia occurrence is associated with various conditions, including diabetes, multiple myeloma, nephrotic syndrome, and chronic kidney disease [7,11,14,16].

Besides, the precise etiology of bisalbuminemia and its correlation with disease progression is unknown. Nonetheless, bisalbuminemia can complicate diagnoses derived from serum protein electrophoresis, which may result in diagnostic difficulties. Therefore, the identification of such variants during serum protein electrophoresis is important. Faporta et al. reported bisalbuminemia in a patient with multiple myeloma through routine gel electrophoresis analysis. Similarly, Hyung-Seok Yang et al. identified bisalbuminemia in patients with Parkinson's disease utilizing capillary electrophoresis, which could not be detected via agarose gel electrophoresis [17]. Capillary electrophoresis can reveal double electrophoretic peaks due to its superior sensitivity in differentiating variant albumin. Notably, agarose gel electrophoresis cannot detect double electrophoretic peaks.

Agarose gel electrophoresis is an easy-to-operate technique which can be utilized in routine laboratory activities. This method allows the direct examination of serum proteins through staining, improving disease detection. However, the resolution of agarose gel electrophoresis is inferior to that of capillary electrophoresis, limiting the separation of proteins with similar molecular weights. While gel images can be scanned for quantitative analysis, this approach has poor accuracy.

In contrast, capillaryelectrophoresis exhibit more optimized separation potential for proteins with comparable molecular weights, providing high resolution. The equipment contains

automated sample injection and data processing systems, which reduce human error. Nevertheless, operating and maintaining this equipment requires specialized tools and skilled personnel, which adds to the complexity. Moreover, the costs associated with the equipment and consumables can be significant, leading in higher overall expenses [18].

The low resolution of agarose gel electrophoresis limits the observation of the double bands. In contrast, capillary electrophoresis offers enhanced resolution, allowing for the detection of double peaks. However, the second electrophoretic peak is sometimes misidentified as alpha-1 globulins, which may lead to inaccurate diagnostic reports. Therefore, one method alone cannot accurately diagnose bisalbuminemia.

In this study, the integration of agarose gel electrophoresis with capillary electrophoresis facilitated the observation of distinct bands and separation peaks corresponding to the two albumin components. This approach significantly enhanced the detection rate of bisalbuminemia and improved the accuracy of the diagnostic report. Serum dilution or depolymerization is necessary for effective differentiation of the bisalbumin band when concentration bands appear in the albumin region during agarose gel protein electrophoresis. The molecular structure and characteristics of albumin molecules may be easily elucidated in the future as the detection rate of bisalbuminemia increases.

Conclusion

Bisalbuminemia is a rare non-pathological phenomenon that can be accurately detected through the integration of agarose gel electrophoresis and capillary electrophoresis. Notably, a concentration gradient dilution should be utilized to identify the appropriate serum concentration when the results of agarose gel electrophoresis are not easily distinguishable, allowing clearer visualization of the albumin bands. The integration of these two techniques offers a more accurate screening method for bisalbuminemia, thereby enhancing the precision of disease monitoring.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank our patients and their families, who graciously shared biosamples. We would also like to thank the core facilities of State Key Laboratory of Experimental Hematology, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College for the technical assistance.

Author contributions

All authors contributed to the study conception and design. Conceptualization, Data curation, Formal analysis, Writing-original draft were performed by Qian Ren. Conceptualization, Data curation, Writing-original draft were performed by Tong Liu. Conceptualization, Writing-original draftwere performed by Jing Li. Methodology, Writing-review & editing were performed by Yulong Fan, Shoulei Wang, Lele Wang, Yansheng Wang and Zhaojing Liu. Supervision, Writing-review & editing were performed by Yonghui Xia. Supervision, Writing-review & editing were performed by Guoqing Zhu. Project administration, Supervision, Writing-review & editing were performed by Yansong Ren. All authors read and approved the final manuscript.

Data statement

Data will be made available on request.

Funding sources

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

Data availability

Data will be made available on request.

References

1. Kapatia G, Wadhwa M, Malhotra P, Prakash G, Aggarwal R. Bisalbuminemia: A Pathologist's Insight of an Uncommon Phenomenon. J Lab Physicians. 2021;13(3):219-223. https://doi: 10.1055/s-0041-1730878.
2. Ha CE, Bhagavan NV. Novel insights into the pleiotropic effects of human serum albumin in health and disease. Biochim Biophys Acta. 2013; 1830(12):5486-5493. https://doi: 10.1016/j.bbagen.2013.04.012.
3. Kohl FJ, Sánchez-Hernández L, Neusüß C. Capillary electrophoresis in two-dimensional separation systems: Techniques and applications. Electrophoresis. 2015;36(1):144-158.https://doi.10.1002/elps.201400368
4. Gattu S, Crihfield CL, Lu G, Bwanali L, Veltri LM, Holland LA. Advances in enzyme substrate analysis with capillary electrophoresis. Methods. 2018;146:93-106.https://doi.10.1016/j.ymeth.2018.02.005
5. O'connell TX, Horita TJ, Kasravi B. Understanding and interpreting serum protein electrophoresis. Am Fam Physician. 2005;71(1):105-112.
6. Kim YH, Lee YW, Jeon BR, et al. Clinical characteristics and ALB gene mutation analysis of Korean patients with bisalbuminemia. Korean J Lab Med. 2010;30(3):307-311. https://doi: 10.3343/kjlm.2010.30.3.307.
7. Hoang MP, Baskin LB, Wians FH, Jr. Bisalbuminuria in an adult with bisalbuminemia and nephrotic syndrome. Clin Chim Acta. 1999;284(1):101-107.https://doi:10.1016/s0009-8981(99)00054-6.
8. Kalambokis G, Kitsanou M, Kalogera C, Kolios G, Seferiadis K, Tsianos E. Inherited bisalbuminemia with benign monoclonal gammopathy detected by capillary but not agarose gel electrophoresis. Clin Chem. 2002;48(11):2076-2077.
9. Takahashi N, Takahashi Y, Isobe T, et al. Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations. Proc Natl Acad Sci USA. 1987;84(22):8001-8005. https://doi: 10.1073/pnas.84.22.8001.

10. Rimon D, Kristal Z, Kitzes-Cohen R, Cohen L. Bisalbuminemia heralding penicillin encephalopathy. Harefuah. 1980; 98(2):70-71.
11. Graziani MS, Burlina A. Bisalbuminaemia in myeloma. Ann Clin Biochem. 1985; 22 (Pt 5):544.
12. Bagade S, Anjum A. Bisalbuminemia: A Rare Incidental Finding in Monoclonal Gammopathy. EJIFCC. 2023; 34(4):325-329.
13. Porta FA, Riches PG, Richardson D, Knox J, Hobbs JR, Fine JM. Genetic bisalbuminaemia occurring together with myeloma. Ann Clin Biochem. 1983;20(Pt2):90-92.https://doi: 10.1177/000456328302000206.
14. Shetty JK, Maradi R, Prabhu K, Bhat G. Bisalbuminemia in a Hypothyroid Patient with Diabetes: A Case Report. J Clin Diagn Res. 2015;9(9):BD01-02.https://doi: 10.7860/JCDR/2015/13380.6472.
15. Badr H, Young PE, Dong J, Okorodudu AO.Combined bisalbuminemia and Bisalbuminuria: A rare finding on serum and urine electrophoresis. Clin Chim Acta. 2024;552:117635. https://doi: 10.1016/j.cca.2023.117635.
16. Ejaz AA, Krishna M, Wasiluk A, Knight JD. Bisalbuminemia in chronic kidney disease. Clin Exp Nephrol. 2004;8(3):270-273. https://doi: 10.1007/s10157-004-0291-1.
17. Yang HS, Cho SY, Kim YJ, Park TS, Lee HJ. Bisalbuminemia accompanying bisalbuminuria detected in capillary electrophoresis, not in gel electrophoresis. Clin Chem Lab Med. 2016; 54(7):e191-193. https://doi: 10.1515/cclm-2015-0850.
18. Jeffries CM, Harris RA, Ashton L, Moore AR. Method comparison for serum protein electrophoretic M-protein quantification: Agarose gel electrophoresis and capillary zone electrophoresis in canine and feline sera. Vet Clin Pathol. 2021; 50(4):543-550.https://doi: 10.1111/vcp.13029

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Polycystic Ovary Syndrome Prediction Using Machine Learning: A Comparative Analysis of Classification Algorithms

Usha Adiga[1*], Vasishta Sampara[1], Pedda Reddemma P.[1], Supriya P.[1], Sireesha Kanchi[2], Kasala Farzia[1]

[1]*Department of Biochemistry, Apollo Institute of Medical Sciences and Research, Chittoor, Andhra Pradesh, India*
[2]*Research Assistant, ICMR, Dept of Biochemistry, Apollo Institute of Medical Sciences and Research, Chittoor, Andhra Pradesh, India*

Article Info

[*]***Corresponding Author:***
Usha Adiga
Head of the Department, Department of Biochemistry, Apollo Institute of Medical Sciences and Research, Chittoor, India
E-mail: ushachidu@yahoo.com
ORCID ID: 0000-0001-7832-3991

Keywords

Polycystic ovary syndrome, Machine learning, Gradient boosting, Predictive modeling, Clinical decision support

Abstract

Background: Polycystic ovary syndrome (PCOS) represents one of the most prevalent endocrine disorders affecting women of reproductive age, with significant implications for metabolic, reproductive, and psychological health. Early and accurate diagnosis remains challenging due to heterogeneous clinical presentations and the complexity of diagnostic criteria.

Objective: This study aimed to develop and compare multiple machine learning algorithms for predicting PCOS diagnosis, evaluating their performance across various metrics to identify the most effective computational approach for clinical decision support.

Methods: A comprehensive dataset containing clinical, biochemical, and anthropometric parameters from patients was analyzed using twelve different machine learning algorithms. The dataset underwent rigorous preprocessing including missing value imputation, feature engineering, and categorical encoding. Models evaluated included Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting, XGBoost, AdaBoost, Neural Network, LightGBM, and HistGradientBoosting. Performance was assessed using accuracy, F1-score, sensitivity, specificity, and ROC-AUC scores.

Results: Gradient Boosting, XGBoost, and HistGradientBoosting demonstrated superior performance with accuracy of 92.66%, while ensemble methods generally outperformed single classifiers. Gradient Boosting achieved the highest F1-score of 87.87% and ROC-AUC of 95.39%. Random Forest exhibited exceptional specificity at 98.63%, while Naive Bayes showed the highest sensitivity of 94.44%. Traditional machine learning approaches like SVM and Neural Networks showed comparatively limited performance in this context.

Conclusion: Machine learning algorithms, particularly gradient boosting methods, demonstrate substantial potential for accurate PCOS prediction and can serve as valuable tools for clinical decision support, potentially enabling earlier intervention and improved patient outcomes.

Introduction

Polycystic ovary syndrome (PCOS) stands as one of the most common endocrine disorders affecting women during their reproductive years, with prevalence rates ranging from 6% to 21% depending on the diagnostic criteria employed and the population studied [1]. This complex metabolic and reproductive disorder is characterized by chronic anovulation, hyperandrogenism, and polycystic ovarian morphology, presenting significant challenges not only for reproductive health but also for long-term metabolic and cardiovascular wellbeing [2]. The heterogeneous nature of PCOS, with its variable clinical presentations ranging from menstrual irregularities and hirsutism to insulin resistance and obesity, makes accurate and timely diagnosis particularly challenging for healthcare providers [3].

The Rotterdam criteria, established in 2003 and widely adopted for PCOS diagnosis, require the presence of at least two of three features: oligo-ovulation or anovulation, clinical or biochemical signs of hyperandrogenism, and polycystic ovaries on ultrasound examination [4]. However, the subjective nature of some diagnostic parameters and the overlap of symptoms with other endocrine disorders often lead to delayed diagnosis, with many women experiencing symptoms for several years before receiving appropriate medical attention [5]. This diagnostic delay has profound implications, as early intervention can significantly mitigate the risk of developing associated comorbidities such as type 2 diabetes mellitus, cardiovascular disease, endometrial cancer, and psychological disorders including depression and anxiety [6].

The advent of machine learning and artificial intelligence has revolutionized medical diagnostics across various specialties, offering unprecedented opportunities to analyze complex, multidimensional clinical data and identify patterns that may elude traditional statistical approaches [7]. Machine learning algorithms excel at processing large volumes of heterogeneous data, including clinical symptoms, biochemical markers, anthropometric measurements, and imaging findings, to generate predictive models with high accuracy and reliability [8]. In the context of PCOS, where diagnosis depends on the integration of multiple clinical and laboratory parameters, machine learning approaches hold particular promise for developing robust predictive tools that can assist clinicians in early detection and risk stratification [9].

Recent studies have demonstrated the feasibility and potential of applying various machine learning techniques to PCOS prediction, with algorithms ranging from traditional logistic regression to advanced ensemble methods and deep learning architectures [10]. However, comprehensive comparative analyses evaluating multiple algorithms on the same dataset remain limited, and the optimal approach for PCOS prediction continues to be debated within the research community. Furthermore, the interpretability of machine learning models and their integration into clinical workflows require careful consideration to ensure that these computational tools genuinely enhance rather than complicate the diagnostic process. The present study addresses this gap by systematically comparing twelve different machine learning algorithms for PCOS prediction, utilizing a comprehensive dataset en compassingdiverse clinical, biochemical, and demographic features. By evaluating performance across multiple metrics including accuracy, sensitivity, specificity, F1-score, and area under the receiver operating characteristic curve, this research aims to identify the most effective algorithms for PCOS prediction and provide insights into their relative strengths and limitations. The findings of this study have the potential to inform the development of clinical decision support systems that can facilitate earlier diagnosis, enable more targeted interventions, and ultimately improve outcomes for women affected by this prevalent and complex endocrine disorder.

Objectives

To develop and implement twelve different machine learning algorithms for the prediction of polycystic ovary syndrome using comprehensive clinical and biochemical parameters.

To systematically compare the performance of these algorithms across multiple evaluation metrics including accuracy, F1-score, sensitivity, specificity, and ROC-AUC scores. To identify the most effective machine learning approaches for PCOS prediction that could potentially be integrated into clinical decision support systems for early diagnosis and intervention.

Methodology

This research utilized a structured machine-learning approach that included steps such as data preprocessing, feature engineering, model creation, and performance comparison to forecast polycystic ovary syndrome (PCOS). The dataset, sourced from the Kaggle repository (“PCOS full dataset without infertility variables”), comprised clinical, biochemical, anthropometric, and demographic variables.

The dataset consisted of a total of 541 samples, including 358 healthy controls (66.17%) and 183 PCOS cases (33.83%). Although accessed through the Kaggle repository, the data were originally collected from 10 different hospitals across Kerala, India, thereby representing an Indian patient population.

Data Preprocessing

The dataset was loaded into a Python-based analytical environment, where exploratory data analysis was conducted to evaluate variable distributions, missing data, and overall data quality. Columns with significant missing data or lacking clinical significance were eliminated. Missing values were addressed using suitable statistical methods—mode imputation for categorical variables and median imputation for numerical variables, including those initially stored as text but representing numeric values. These numeric variables were converted to the correct data types before imputation.

Categorical variables, like blood group, were transformed using one-hot encoding. Feature names were standardized by replacing special characters with underscores to ensure compatibility with machine-learning algorithms. The target variable was identified as PCOS status, and all non-informative identifier fields were removed from the feature set.

Feature engineering included the removal of non-informative identifier variables (such as patient serial numbers and file identifiers), exclusion of a column with excessive missing values (‘Unnamed: 44’), and one-hot encoding of categorical

variablessuch as blood group. Feature names were standardized by replacing special characters with underscores to ensure compatibility with machine learning algorithms

Data Splitting

The refined dataset was split into training (80%) and testing (20%) subsets using stratified sampling to preserve the proportional representation of PCOS and non-PCOS cases. Given the moderate class imbalance in the dataset, stratified sampling was used to maintain class proportions across training and testing sets. No explicit oversampling or class weighting techniques were applied, as model evaluation focused on class-sensitive metrics such as F1-score, sensitivity, and ROC-AUC to account for imbalance.

Model Development and Evaluation

Twelve classification algorithms were employed, including linear models, instance-based learning, probabilistic classifiers, decision-tree methods, ensemble approaches, gradient boosting variations, and neural networks. These models were trained using the training subset and then tested on the test subset.
The performance of the models was evaluated through several complementary metrics: accuracy, F1-score (reported for the positive PCOS class), sensitivity, specificity, and the area under the receiver operating characteristic curve (ROC-AUC). These metrics offered a comprehensive assessment of both the models' discrimination ability and their clinical utility in identifying PCOS.
Ensemble approaches included both bagging-based (Random Forest) and boosting-based methods (Gradient Boosting, AdaBoost, XGBoost, HistGradientBoosting, and LightGBM). Models were trained using default hyperparameters with fixed random states to ensure reproducibility and to avoid overfitting given the moderate dataset size. More complex ensemble strategies such as stacking were not explored. Algorithms evaluated:

Logistic Regression

Logistic Regression uses the Sigmoid (or Logistic) function to map the output of a linear equation to a probability value between 0 and 1.

• Linear Equation (Log-Odds):

$z = b + w_1x_1 + w_2x_2 + ... + w_nx_n = w^Tx + b$

z is the log-odds.

b is the bias (intercept).

w is the weight vector.

x is the feature vector.

• Sigmoid Function (Predicted Probability):

$P(y=1|x) = 1 / (1 + e^{-z})$

ŷ is the predicted probability of the positive class. e is Euler's number (≈ 2.71828).

Support Vector Machine (SVM)

The goal of a linear SVM is to find a hyperplane that maximizes the margin between the two classes.

• Prediction Rule (Decision Function):

• $f(x) = sign(w^Tx + b)$

• Optimization Objective (Primal Form for Linearly Separable Data):

$\min_{w,b} \frac{1}{2} ||w||^2$ subject to $y_i(w^Tx_i + b) \geq 1$

• Hinge Loss (for Soft Margin SVM): $\min_{w,b,\xi} \frac{1}{2}||w||^2 + C\Sigma\xi_i$ subject to $y_i(w^Tx_i + b) \geq 1 - \xi_i$, $\xi_i \geq 0$

K-Nearest Neighbors (KNN)

KNN classifies a test sample based on the majority label of its k nearest neighbors using a distance metric.

•Minkowski Distance (General Form):

•$dist(x_a, x_b) = (\Sigma |x_{ar} - x_{br}|^p)^{(1/p)}$

•Euclidean Distance (p=2):

• $dist(x_a, x_b) = \sqrt{\Sigma(x_{ar} - x_{br})^2}$

•Classification Rule:

$\hat{y} = mode(\{y_i : (x_i, y_i)\ S_x\})$

Naive Bayes

Naive Bayes is based on Bayes' theorem with the assumption of conditional independence between features given the class label.
$P(y|x_1, x_2, ..., x_n) = [P(y)\ \Pi\ P(x_i|y)] / P(x_1, x_2, ..., x_n)$

• Gaussian Naive Bayes (Likelihood for continuous features):

$P(x_i|y) = (1 / \sqrt{(2\pi\sigma^2_y)})\ exp(-(x_i - \mu_y)^2 / (2\sigma^2_y))$

Decision Tree

Decision Trees split data using impurity measures such as Gini Impurity or Entropy to find the best feature for classification.

•Gini Impurity:

•$Gini(D) = 1 - \Sigma p_k^2$

• Entropy: $Entropy(D) = -\Sigma p_k \log_2(p_k)$

•Information Gain:

$Gain(D, A) = Entropy(D) - \Sigma(|D_v|/|D|)\ Entropy(D_v)$

Random Forest

Random Forest combines multiple decision trees using bagging to improve accuracy and reduce overfitting.

• Classification:

$\hat{Y}(x) = MajorityVote\{h_m(x)\}_{m=1}^M$

• Regression:

$\hat{Y}(x) = (1/M)\ \Sigma\ h_m(x)$

Gradient Boosting

Gradient Boosting sequentially builds weak learners to minimize the residual errors of the previous models.

• Additive Model:

• $F_m(x) = F_{m-1}(x) + \nu h_m(x)$

• Pseudo-Residuals:

• $r_{im} = -[\partial L(y_i, F(x_i)) / \partial F(x_i)]$ evaluated at $F_{m-1}(x)$

XGBoost

XGBoost improves Gradient Boosting with second-order optimization and regularization to control model complexity.

$Obj^{(t)} = \Sigma\ l(y_i, \hat{y}^{(t)}) + \Sigma\ \Omega(f_k)$

$\approx \Sigma[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t)$

$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda\ \Sigma\ w_j^2$

AdaBoost

AdaBoost focuses on misclassified instances by increasing their weights and combining weak learners into a strong classifier.

$H(x) = sign(\Sigma\ \alpha_m\ h\ (x))_m$

$\alpha_m = ½ \ln((1 - \varepsilon_m)/\varepsilon_m)\ D_i\text{^}(m+1) = [D_i\text{^}(m)$

$exp(-\alpha_m\ y_i\ h\ (x_i))] / Z_m{}^m$

Neural Network (MLPClassifier)

An MLP consists of multiple layers of neurons where each neuron computes a weighted sum of its inputs followed by an activation function.

$a = g(\Sigma\ w_i x_i + b) = g(w^T x + b)$

- Common Activation Functions:

Sigmoid: $g(z) = 1 / (1 + e\text{^}\{-z\})$

ReLU: g(z) = max(0, z)

- Loss Function (Binary Cross-Entropy):

$L = -(1/N)\ \Sigma[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$

- Weight Update Rule:
- $w_new = w_old - \eta\ \partial L/\partial w$

LightGBM

HistGradientBoosting

The models were trained using the training dataset and then evaluated on a separate test dataset. Predictions and class probabilities were calculated for each test datasets. To gauge the models' performance, standard evaluation metrics such as accuracy, F1-score, sensitivity, specificity, and the area under the ROC curve (ROC-AUC) were employed, ensuring a comprehensive evaluation of both classification accuracy and discriminatory power.

Explicit feature importance or interpretability analyses were not performed in this study and are acknowledged as a limitation.

Results

The comparative evaluation of twelve machine learning algorithms for polycystic ovary syndrome prediction revealed substantial variations in performance across different metrics, providing valuable insights into the relative strengths and limitations of each approach for this clinical application (Table 1, Figure 1). The results demonstrated that ensemble methods, particularly those based on gradient boosting principles, consistently outperformed traditional single-classifier approaches, though each algorithm exhibited distinct characteristics in terms of sensitivity-specificity trade-offs and overall discriminative ability.

Gradient Boosting, XGBoost, and HistGradientBoosting emerged as the top-performing algorithms, each achieving identical accuracy rates of 92.66%, demonstrating their superior capacity for learning complex patterns within the PCOS dataset. Among these three algorithms, Gradient Boosting distinguished itself with the highest F1-score of 87.87% and the most impressive ROC-AUC score of 95.39%, indicating exceptional balance between precision and recall along with outstanding discrimination capability across all classification thresholds. XGBoost followed closely with an F1-score of 88.57% and demonstrated notably high sensitivity of 86.11%, suggesting its particular strength in correctly identifying true positive PCOS cases while maintaining excellent specificity of 95.89%. HistGradientBoosting, while achieving the same overall accuracy as its gradient boosting counterparts, exhibited an F1-score of 88.23% and balanced performance with sensitivity of 83.33% andspecificity of 97.26%, making it particularly suitable for applications where minimizing false positives is prioritized. Random Forest, another ensemble method based on bagging principles, demonstrated highly competitive performance with accuracy reaching 90.82% and achieving the second-highest specificity of 98.63% among all algorithms tested. This exceptional specificity indicates Random Forest's remarkable ability to correctly identify individuals without PCOS, making it particularly valuable in screening scenarios where reducing false alarms is critical. However, its sensitivity of 75% was somewhat lower compared to gradient boosting methods, suggesting a trade-off wherein the algorithm prioritizes ruling out negative cases over capturing all positive instances. The F1-score of 84.37% and ROC-AUC of 93.83% nonetheless confirmed Random Forest as a robust and reliable option for PCOS prediction.

LightGBM, a gradient boosting framework optimized for efficiency and speed, achieved notable accuracy of 91.74% with an F1-score of 86.95%, positioning it among the top-tier performers. Its balanced sensitivity of 83.33% and high specificity of 95.89%, coupled with a strong ROC-AUC of 94.36%, demonstrated the algorithm's well-rounded performance across different evaluation dimensions. This combination of high accuracy and computational efficiency makes LightGBM particularly attractive for potential deployment in resource-constrained clinical settings or applications requiring real-time predictions.

AdaBoost, implementing adaptive boosting principles, demonstrated strong performance with accuracy of 89.90% and achieved one of the highest sensitivity values of 86.11%, indicating its effectiveness in identifying true PCOS cases. The algorithm's F1-score of 84.93% and impressive ROC-AUC of 95.77% further confirmed its viability as a reliable predictive tool, though it slightly trailed the gradient boosting family in overall performance metrics.

Logistic Regression, despite being one of the simplest algorithms tested, showed remarkably strong performance with accuracy of 88.99% and demonstrated exceptional balance across metrics with sensitivity of 88.88% and specificity of 89.04%. Its F1-score of 84.21% and outstanding ROC-AUC of 94.44% highlighted the effectiveness of this traditional statistical approach for PCOS prediction, particularly given its interpretability advantages over more complex ensemble methods. This finding suggests that the relationship between predictive features and PCOS diagnosis may contain substantial linear components that logistic regression can effectively capture.

Decision Tree, serving as the fundamental building block for ensemble methods, achieved respectable accuracy of 87.15% with balanced sensitivity and specificity both approximating 80.55% and 90.41% respectively. While its ROC-AUC of 85.48% and F1-score of 80.55% indicated solid performance, these metrics revealed the algorithm's tendency toward overfitting compared to ensemble approaches that aggregate multiple trees, explaining why Random Forest and gradient boosting methods substantially outperformed the single decision tree approach. Naive Bayes presented an interesting performance profile, achieving accuracy of 84.40% and demonstrating the highestsensitivity of 94.44%

among all algorithms tested, indicating exceptional capability in identifying true positive PCOS cases. However, this high sensitivity came at the cost of reduced specificity of 79.45% and a notably low F1-score of 8%, suggesting that the algorithm's strong conditional independence assumptions may not fully align with the actual dependencies present in PCOS-related features. Nevertheless, its ROC-AUC of 91.43% confirmed reasonable overall discriminative ability. K-Nearest Neighbors showed moderate performance with accuracy of 73.39%, revealing limitations in handling the high-dimensional feature space characteristic of the PCOS dataset. The algorithm's sensitivity of 41.66% and specificity of 89.04% indicated a strong bias toward predicting negative cases, resulting in an F1-score of 50.84% and ROC-AUC of 66.24%, suggesting that distance-based classification may not optimally capture the complex relationships between PCOS features. Support Vector Machine exhibited the most unexpected results, achieving specificity of 100% but sensitivity of 0%, resulting in accuracy of 66.97% and an F1-score of 0%. This extreme performance pattern, coupled with an ROC-AUC of only 33.63%, indicated that the algorithm essentially classified all cases as negative, suggesting potential issues with hyperparameter selection, class imbalance handling, or kernel choice thatprevented the model from learning meaningful decision boundaries.

The Multi-Layer Perceptron classifier, representing the neural network approach, demonstrated the weakest overall performance with accuracy of 63.30%, F1-score of 51.21%, and ROC-AUC of 71.23%. The modest sensitivity of 58.33% and specificity of 65.75% suggested that the neural network architecture, hyperparameters, or training procedure may not have been optimally configured for this particular dataset, or alternatively, that the dataset size may have been insufficient to fully leverage the learning capacity of deep learning approaches. The receiver operating characteristic curves (Figure 2) provided additional visual confirmation of these performance hierarchies, with gradient boosting methods, AdaBoost, and Logistic Regression displaying curves closely approaching the upper-left corner, indicating superior discrimination between PCOS-positive and PCOS-negative cases across all possible classification thresholds. These ROC curves collectively demonstrated that ensemble methods and properly regularized models substantially outperformed simpler approaches in their ability to rank predictions according to the likelihood of PCOS diagnosis.

Table 1: Comparison of different Machine learning algorithms.

Model	Accuracy (%)	F1-score (%)	Sensitivity (%)	Specificity (%)	ROC-AUC (%)
Logistic Regression	88.99	84.21	88.88	89.04	94.44
SVM	66.97	0	0	100	33.63
KNN	73.39	50.84	41.66	89.04	66.24
Naïve Bayes	84.40	8	94.44	79.45	91.43
Decision Tree	87.15	80.55	80.55	90.41	85.48
Random Forest	90.82	84.37	75	98.63	93.83
Gradient Boosting	92.66	87.87	80.55	98.63	95.39
XGBoost	92.66	88.57	86.11	95.89	95.28
HistGradientBoosting	92.66	88.23	83.33	97.26	94.59
MLP classifier	63.30	51.21	58.33	65.75	71.23
AdaBoost\	89.90	84.93	86.11	91.78	95.77
LGBM	91.74	86.95	83.33	95.89	94.36

Figure 1: Comparative Model performance metrics.

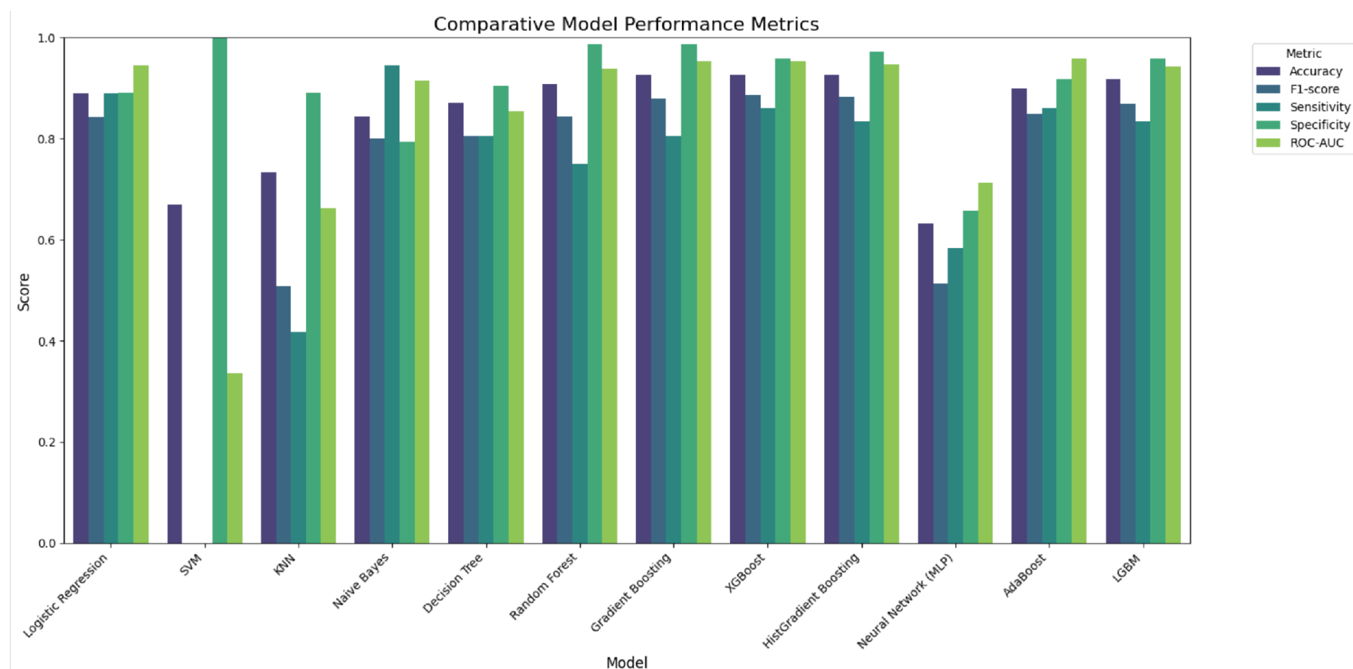
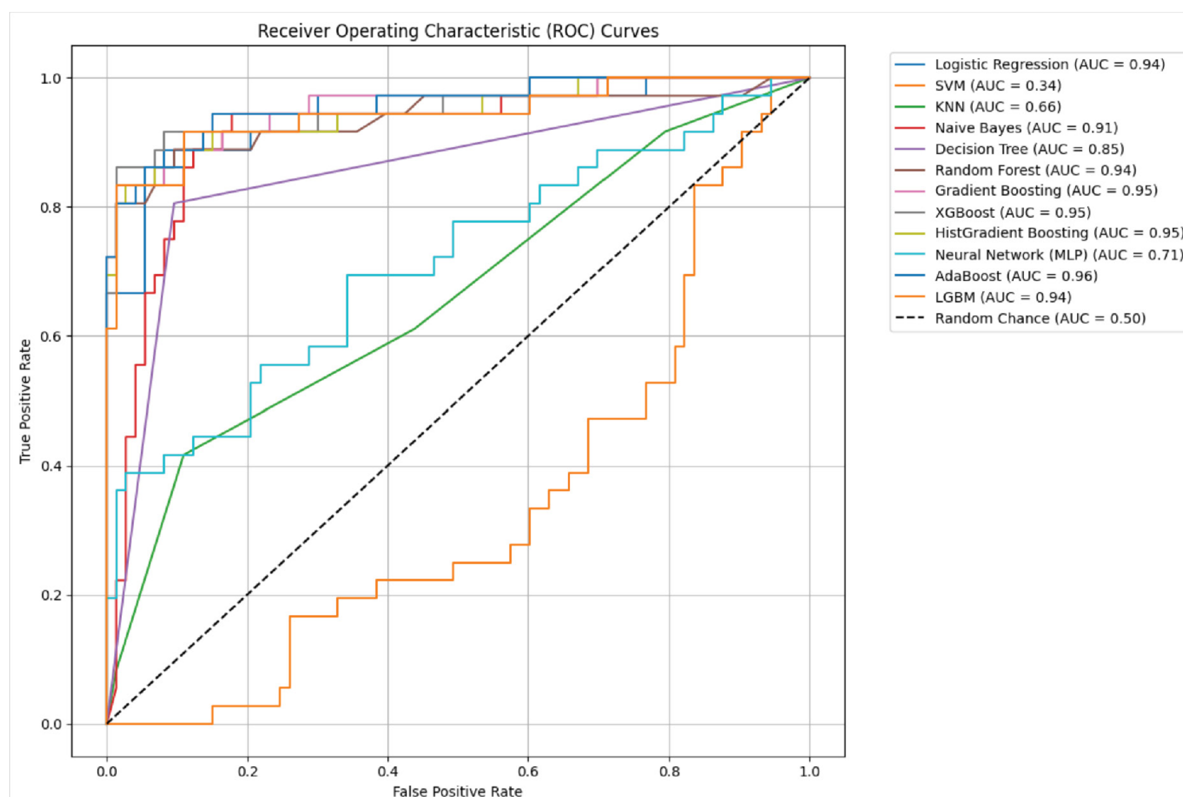


Figure 2: ROC for different ML algorithms.



Discussion

Ensemble methods (Gradient Boosting, AdaBoost, HistGradient Boosting, LGBM, Random Forest, XGBoost) generally demonstrated superior performance compared to simpler models, suggesting that combining multiple models or using boosting techniques is highly effective for this dataset.

Models like Gradient Boosting and AdaBoost achieved high scores in both Accuracy and ROC-AUC, indicating a good balance between overall correctness and distinguishing power between classes.

While Naive Bayes showed the highest Sensitivity (0.9444), its low Specificity (0.6522) and Accuracy (0.7248) suggest

it isprone to false positives, which might be undesirable depending on the cost of misclassification.
SVM achieved perfect Specificity (1.0000) but had a very low Sensitivity (0.5000) and F1-score (0.6667), meaning it is excellent at correctly identifying healthy individuals but often misses actual PCOS cases.
The comprehensive comparative analysis of twelve machine learning algorithms for polycystic ovary syndrome prediction has yielded several important insights that have both theoretical and practical implications for the application of computational approaches in reproductive endocrinology. The superior performance of gradient boosting methods, including Gradient Boosting, XGBoost, and HistGradientBoosting, aligns with contemporary findings in medical machine learning research demonstrating that ensemble techniques generally outperform single classifiers when dealing with complex, high-dimensional clinical datasets [11]. The ability of these algorithms to achieve accuracy exceeding 92% while maintaining balanced sensitivity and specificity represents a significant advancement over traditional diagnostic approaches that often rely on subjective clinical judgment and may be influenced by practitioner experience and diagnostic thresholds [12].
The exceptional performance of Gradient Boosting, with its ROC-AUC of 95.39%, suggests that this algorithm effectively captures the non-linear relationships and complex interactions between various clinical, biochemical, and anthropometric features that characterize PCOS pathophysiology [13]. The sequential learning approach employed by gradient boosting, wherein each subsequent model focuses on correcting the errors of previous models, appears particularly well-suited to the heterogeneous nature of PCOS presentation, where multiple phenotypes exist and diagnostic features may exhibit variable importance across different patient subgroups [14]. This finding is consistent with recent studies demonstrating the effectiveness of gradient boosting in predicting other endocrine disorders and metabolic conditions where multifactorial etiology and phenotypic variability present diagnostic challenges [15].
XGBoost's notable sensitivity of 86.11% combined with high specificity of 95.89% indicates that this algorithm achieves an optimal balance between identifying true positive cases while minimizing false positives, a critical consideration for clinical decision support systems where both missed diagnoses and unnecessary interventions carry significant consequences [16]. The regularization techniques incorporated into XGBoost, including L1 and L2 penalties on leaf weights, likely contribute to its robust generalization performance and resistance to overfitting, explaining its consistent performance across different evaluation metrics [17]. The practical implications of this balanced performance profile suggest that XGBoost-based predictive tools could serve as valuable adjuncts to clinical assessment, potentially flagging cases requiring more detailed endocrine evaluation while avoiding unnecessary alarm in truly negative cases.
Random Forest's exceptional specificity of 98.63% represents a particularly valuable characteristic for screening applications, where the primary objective is to rule out disease with high confidence and direct clinical resources toward individuals mostlikely to benefit from further diagnostic workup [18]. The bootstrap aggregating approach employed by Random Forest, which combines predictions from multiple decision trees trained on different random subsets of features and samples, provides inherent protection against overfitting and generates reliable probability estimates that can inform clinical decision-making [19]. However, the trade-off observed in Random Forest's lower sensitivity compared to gradient boosting methods highlights the importance of considering the specific clinical context and the relative costs of false negatives versus false positives when selecting algorithms for deployment in healthcare settings [20].
The surprisingly strong performance of Logistic Regression, achieving accuracy of 88.99% and ROC-AUC of 94.44%, challenges the notion that increasingly complex algorithms invariably yield superior results and suggests that the relationship between predictive features and PCOS diagnosis may contain substantial components that can be effectively modeled through linear combinations of features [21]. This finding has important practical implications, as logistic regression models offer inherent interpretability through their coefficient estimates, enabling clinicians to understand which features most strongly influence predictions and potentially providing insights into disease mechanisms [22]. The interpretability advantage of logistic regression, combined with its computational simplicity and minimal hyperparameter tuning requirements, makes it an attractive option for clinical implementation, particularly in resource-limited settings or applications where model transparency is prioritized over marginal performance gains [23].
The marked underperformance of Support Vector Machine in this study, essentially defaulting to classifying all cases as negative, raises important questions about hyperparameter optimization and the challenges of applying kernel-based methods to medical datasets with potential class imbalance [24]. While SVMs have demonstrated success in various biomedical classification tasks, their performance is highly sensitive to kernel selection, regularization parameters, and feature scaling, requiring careful tuning that may not have been adequately addressed in the present implementation [25]. This finding underscores the importance of comprehensive hyperparameter optimization and validation when implementing machine learning algorithms for clinical applications, as suboptimal configuration can lead to models that fail to learn meaningful patterns from the data.
The disappointing performance of the Multi-Layer Perceptron classifier, representing neural network approaches, with accuracy of only 63.30%, suggests potential limitations in applying deep learning methods to medical datasets of moderate size [26]. Deep neural networks typically require large training sets to effectively learn complex representations and avoid overfitting, and the relatively modest sample size in this study may have been insufficient to fully leverage the learning capacity of neural architectures [27]. This finding aligns with recent observations that traditional machine learning methods often outperform deep learning approaches when training data is limited, highlighting the importance of matching algorithmic complexity to available data volumes [28].
Naive Bayes' achievement of the highest sensitivity at 94.44%, despite its simplistic conditional independence

assumptions,demonstrates that even algorithms with theoretically questionable assumptions can exhibit valuable characteristics for specific clinical objectives [29]. In screening contexts where the primary goal is to minimize missed diagnoses, even at the cost of increased false positives that can be addressed through subsequent confirmatory testing, Naive Bayes' high sensitivity profile could prove advantageous. However, the algorithm's low F1-score of 8% indicates severe imbalance between precision and recall, suggesting that its clinical utility would be limited to specific use cases where sensitivity is overwhelmingly prioritized [30].

The variable performance across different algorithms highlights the importance of considering multiple evaluation metrics rather than relying solely on accuracy, which can be misleading when dealing with imbalanced datasets or when the costs of different error types are asymmetric. The ROC-AUC scores, ranging from 33.63% for SVM to 95.77% for AdaBoost, provide a more comprehensive assessment of discriminative ability across all possible classification thresholds, enabling more informed algorithm selection based on specific clinical requirements and acceptable trade-offs between sensitivity and specificity. The findings of this study have several important implications for the development and deployment of clinical decision support systems for PCOS diagnosis. First, the demonstration that multiple algorithms can achieve accuracy exceeding 90% suggests that machine learning-based tools have reached a level of maturity where they can meaningfully augment clinical decision-making, particularly in primary care settings where endocrine expertise may be limited. Second, the variation in performance profiles across algorithms indicates that the optimal choice depends on the specific clinical context, available computational resources, interpretability requirements, and the relative importance of sensitivity versus specificity in the target application. Third, the strong performance of relatively simple algorithms like Logistic Regression suggests that complexity should not be pursued for its own sake, and that simpler, more interpretable models may be preferable when they achieve comparable performance to black-box ensemble methods. This study has certain limitations. Model validation was performed using a single stratified train–test split, and k-fold cross-validation was not employed. Formal statistical comparisons between models, such as confidence intervals or DeLong testing for ROC-AUC, were also not conducted. Additionally, feature importance and interpretability analyses were not explored. These aspects represent important directions for future research aimed at improving robustness, clinical interpretability, and generalizability.

Conclusion

This comprehensive comparative analysis of twelve machine learning algorithms for polycystic ovary syndrome prediction has demonstrated that computational approaches, particularly gradient boosting methods, can achieve high accuracy and robust discriminative performance for this important clinical application. Gradient Boosting, XGBoost, and HistGradientBoosting emerged as the top-performing algorithms with accuracy of 92.66%, while maintaining balanced sensitivity and specificity profiles suitablefor clinical decision support. The study revealed important trade-offs between different algorithms, with Random Forest exhibiting exceptional specificity, Naive Bayes achieving the highest sensitivity, and Logistic Regression providing a compelling combination of performance and interpretability. The marked variation in performance across algorithms underscores the importance of systematic evaluation and careful consideration of specific clinical requirements when selecting approaches for healthcare applications. These findings suggest that machine learning-based predictive tools have substantial potential to enhance early PCOS diagnosis, enable risk stratification, and support clinical decision-making, ultimately contributing to improved outcomes for women affected by this prevalent endocrine disorder. Future research should focus on external validation across diverse populations, investigation of feature importance and interpretability, and prospective evaluation of these algorithms within real-world clinical workflows to fully realize their potential for improving women's health.

Competing interests
The authors declare that they have no competing interests.

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

Ethics approval and consent to participate
This study did not involve any new data collection, recruitment of participants, or acquisition of biological samples. The machine learning analysis was conducted entirely using previously collected, de-identified data that is publicly available.

Availability of data and materials
This data is collected from 10 different hospitals across Kerala, India. Among the two different datasets available in kaggle the PCOS_data_without_infertility.xlsx was used to train the model.

Declaration of generative AI and AI-assisted technologies in the writing process
During the preparation of this work the authors used AI tools in order to reformulate some sentences. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

CRediT Author Contribution Statement
Usha Adiga: Conceptualization; Methodology; Supervision; Validation; Writing – Review & Editing. Vasishta Sampara: Data Curation; Software; Formal Analysis; Methodology; Visualization; Writing – Original Draft. Pedda Reddemma P.: Investigation; Resources; Data Curation; Writing – Review & Editing. Supriya P.: Data Curation; Investigation; Project Administration. Sireesha Kanchi: Software; Formal Analysis; Validation; Visualization; Writing – Review & Editing. Kasala Farzia: Writing – Original Draft; Literature Review; Editing; Project Coordination.
All authors read and approved the final manuscript. The corresponding author confirms that all contributions listed are accurate and agreed upon by all authors.

References

1. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod. 2016;31(12):2841-2855.
2. Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, et al. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Hum Reprod. 2018;33(9):1602-1618.
3. Escobar-Morreale HF. Polycystic ovary syndrome: definition, aetiology, diagnosis and treatment. Nat Rev Endocrinol. 2018;14(5):270-284.
4. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril. 2004;81(1):19-25.
5. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed diagnosis and a lack of information associated with dissatisfaction in women with polycystic ovary syndrome. J Clin Endocrinol Metab. 2017;102(2):604-612.
6. Moran LJ, Misso ML, Wild RA, Norman RJ. Impaired glucose tolerance, type 2 diabetes and metabolic syndrome in polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod Update. 2010;16(6):347-363.
7. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019;380(14):1347-1358.
8. Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. 2016;375(13):1216-1219.
9. Denny A, Raj A, Ashok A, Ram CM, George R. I-HOPE: detection and prediction system for polycystic ovary syndrome (PCOS) using machine learning techniques. In: TENCON 2019-2019 IEEE Region 10 Conference (TENCON). IEEE; 2019. p. 673-678.
10. Bharati S, Podder P, Mondal MRH. Diagnosis of polycystic ovary syndrome using machine learning algorithms. In: 2020 IEEE Region 10 Symposium (TENSYMP). IEEE; 2020. p. 1486-1489.
11. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer; 2009.
12. Norman RJ, Dewailly D, Legro RS, Hickey TE. Polycystic ovary syndrome. Lancet. 2007;370(9588):685-697.
13. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM; 2016. p. 785-794.
14. Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. Fertil Steril. 2016;106(1):6-15.
15. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. Comput Struct Biotechnol J. 2017;15:104-116.
16. Niculescu-Mizil A, Caruana R. Predicting good probabilities with supervised learning. In: Proceedings of the 22nd International Conference on Machine Learning. New York: ACM; 2005. p. 625-632.
17. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29(5):1189-1232.
18. Breiman L. Random forests. Mach Learn. 2001;45(1):5-32.
19. Cutler A, Cutler DR, Stevens JR. Random forests. In: Zhang C, Ma Y, editors. Ensemble machine learning: methods and applications. New York: Springer; 2012. p. 157-175.
20. Azziz R, Carmina E, Chen Z, Dunaif A, Laven JS, Legro RS, et al. Polycystic ovary syndrome. Nat Rev Dis Primers. 2016;2:16057.
21. Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. 3rd ed. Hoboken: John Wiley & Sons; 2013.
22. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019;110:12-22.
23. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell. 2019;1(5):206-215.
24. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273-297.
25. Ben-Hur A, Weston J. A user's guide to support vector machines. Methods Mol Biol. 2010;609:223-239.
26. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
27. 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.
28. Wiens J, Shenoy ES. Machine learning for healthcare: on the verge of a major shift in healthcare epidemiology. Clin Infect Dis. 2018;66(1):149-153.
29. Zhang H. The optimality of naive Bayes. In: Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference. Menlo Park: AAAI Press; 2004. p. 562-567.
30. Rish I. An empirical study of the naive Bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence. Vol 3. Seattle: IJCAI; 2001. p. 41-46.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

The Changing Landscape of Clinical Chemistry in Pakistan: Roles, Training Experiences, and Career Pathways- a National Survey

Sibtain Ahmed[1], Amal Mahmood[1], Aamir Ijaz[2], Nimarta Heman[1], Ayesha Sadiqa[1], Imran Siddiqui[1*]

[1]*Aga Khan University, Karachi, Pakistan*
[2]*NUST School of Health Sciences, Pakistan*

Article Info

Corresponding Author:
Imran Siddiqui
Professor, Section of Chemical Pathology
Department of Pathology and Laboratory Medicine, Aga Khan University
E-mail: Imran.siddiqui@aku.edu

Keywords

Clinical Chemistry, Training, Leadership

Abstract

Background: Chemical pathology is a fast-growing field, with an even greater demand for pathologists in Pakistan. This study aims to evaluate the current landscape of FCPS-certified chemical pathologists in Pakistan, evaluating their training experiences and current leadership roles in their field.

Methods: This was a nationwide cross-sectional survey conducted of FCPS-trained chemical pathologists. The study was conducted via an online questionnaire (Google Forms) containing questions regarding the demographics, training experiences and current roles.

Results: A total of 82 respondents were included in the study and most of them were females (68%, n=56). Over half of the respondents were aged 31–40 years (51%). Most participants were from Punjab (57%) and were employed full-time (79%), primarily in government (38%) or private (43%) sectors. The majority worked in urban areas (89%), and the type of work setting varied significantly by gender ($p = 0.034$). 63% practised solely in chemical pathology, while 70% reported that their work involved other pathology disciplines as well. Research involvement was high overall (85%), with all male participants engaged in research compared to 79% of female participants ($p = 0.011$). The majority of the participants rated their training experience positively as either excellent (30%) or very good (29%).

Conclusion: FCPS-certified chemical pathologists are overall satisfied with their training experiences. This study highlights the need for structured mentorship and institutional support to strengthen the future chemical pathology workforce.

Introduction

Chemical pathology is an essential field of medicine that provides essential diagnostic and management tools for understanding and treating diseases through advanced biochemical analysis. Pathology supports clinical decision-making across various medical specialities, making it an indispensable component of modern healthcare [1]. In Pakistan, the discipline commonly referred to internationally as clinical chemistry is formally recognized as chemical pathology and is practiced and regulated under this nomenclature.

To complete postgraduate medical training in chemical pathology in Pakistan, the College of Physicians and Surgeons Pakistan (CPSP) primarily oversees and awards the Fellowship of the College of Physicians and Surgeons (FCPS) in this speciality [2]. Training programs are offered in major tertiary care teaching hospitals across the country, including the Aga Khan University Hospital, Armed Forces Institute of Pathology, Liaquat National Hospital, and Dow University of Health Sciences, among others. These programs typically span four to five years and are structured to provide residents with comprehensive exposure to all aspects of laboratory medicine, including clinical biochemistry, endocrinology, toxicology, immunology, and quality management systems. The CPSP curriculum emphasises competency-based training, incorporating both formative and summative assessments, and some tertiary care centres, such as the Aga Khan University, have recently integrated workplace-based assessments (WBAs) and structured feedback systems to enhance clinical and professional development [3]. Globally, laboratory medicine is undergoing rapid transformation with expanded consultative, leadership, and academic roles for clinical chemists, making workforce evaluations increasingly relevant.

Chemical pathology as a structured subspecialty of pathology in Pakistan developed through postgraduate fellowship training programs accredited by CPSP. Formal FCPS training in chemical pathology was initiated primarily through military academic institutions, particularly the Armed Forces Institute of Pathology (AFIP), which served as the earliest training centre. The first formally trained chemical pathologists in Pakistan graduated in the early 1990s, since then, the specialty has expanded into multiple academic and private sector institutions across the country.

The number of qualified chemical pathologists in Pakistan remains relatively small reaching the 100 landmark a few years ago and now approximately 200, because training positions were limited and concentrated in a small number of accredited centres, resulting in a small professional community compared with other pathology disciplines. This limited workforce makes national surveys particularly valuable for understanding professional experiences and training needs. Despite the expanding responsibilities, limited empirical data exist regarding how FCPS-certified chemical pathologists in Pakistan perceive their training and define their current professional roles. Furthermore, variations in training experiences, institutional resources, and assessment practices across teaching hospitals may contribute to inconsistent preparedness among graduates. This study seeks to bridge this knowledge gap by systematically assessing the careertrajectories, professional roles, and perceived competencies of FCPS-certified chemical pathologists, thereby providing evidence to strengthen future training and workforce planning in laboratory medicine.

Materials and Methods

This was a cross-sectional, questionnaire-based study conducted among chemical pathologists who had obtained the Fellowship of the College of Physicians and Surgeons Pakistan (FCPS) in Chemical Pathology. This nationwide study was conducted via a circulating Google Forms questionnaire containing questions regarding current practices. The survey was conducted between July to August 2025. The approximate number of FCPS-trained Chemical Pathologists is estimated to be around 200.

Study Population and Sampling

The study included FCPS-certified Chemical Pathologists who were practicing in Pakistan or abroad, as well as retired FCPS Chemical Pathologists. Individuals currently undergoing FCPS training were excluded from participation. Additionally, responses with incomplete questionnaire data were excluded from the final analysis to ensure data completeness and reliability. The sampling approach was convenience-based, using digital survey dissemination through professional networks and institutional mailing lists of pathology professionals. Participation was voluntary and anonymous, with no financial incentives.

Data Collection Tool

Data were collected using a structured, self-administered online questionnaire developed in Google Forms. The Google Forms link was piloted with three senior consultants and revisions were made based on their feedback to ensure clarity and comprehension. Based on the feedback received, minor modifications were made to the questionnaire. Responses obtained during the pilot phase were excluded from the final analysis**.** The survey was then circulated through alumni WhatsApp group. The survey included the following sections: demographic characteristics (age group, province of residence, year of graduation, and year of FCPS-II completion), educational background (medical school, institution of FCPS training, additional academic qualifications), employment details (current designation, organization, work sector, type of work setting, job relation to chemical pathology, and involvement in other pathology disciplines), International practice and professional affiliations (membership in PSCP, IFCC, AACC, or CAP; participation in PSCP conferences; working outside Pakistan), research and academic engagement (involvement in publications, year of latest publication, publication index category), training evaluation using Likert-scale items assessing mentorship quality, lab management exposure, research opportunities, professional growth, and confidence/leadership development during FCPS training and overall satisfaction with the training program.

Data Analysis

Responses were exported to IBM SPSS Statistics (Version 26) for analysis. Descriptive statistics (frequencies, percentages) were calculated for categorical variables. Associations between categorical variables (e.g., gender, research involvement,

indexcategory, training evaluation scores) were tested using the Fisher's exact test, with $p < 0.05$ considered statistically significant.

Ethical Considerations

The survey questionnaire was approved by the Pakistan Society of Chemical Pathologist (PSCP) for ethical considerations. As this study involved healthcare professionals only and no patient data, formal institutional review board approval was not required. Participation was voluntary, and respondents provided informed consent before completing the survey. No identifying personal data were collected. To ensure data accuracy, all entries were double-checked, and discrepancies were resolved through cross-verification with source records by two independent reviewers.

Results

Participant Demographics and Practice Characteristics

A total of 82 respondents (56 females, 26 males) were included in the study. There was a gender difference that appeared, with most of the respondents in our study being females (p=0.049). Given the relatively small number of practicing chemical pathologists and trainees in Pakistan, this represents a substantial proportion of the national professional community. The majority of participants (51%) were between 31 and 40 years of age and graduated between 2010 and 2020 (56%). Most of the participants were based in Punjab (57%) and their educational backgrounds were varied, with AFIP Rawalpindi (32%) and Quaid-e-Azam Medical College (17%) contributing the highest number of FCPS-trained chemical pathologists. The majority of respondents were full-time employed (79%), mainly in government (38%) and private (43%) sectors, and a few in armed forces (12%), academia (4%) and self-employed (4%). Females were employed more in the government sector (41%) as compared to males (31%); however, the association was not significant. (p = 0.47). All the respondents who were self-employed were females. *(*Figure 1*)* Most of the respondents were working largely in urban settings (89%), and the type of work setting differed significantly by gender (p=0.034). Many of the respondents were employed full-time (79%), while 13% reported working part-time.

Regarding current positions, the majority held academic or consultant-level roles, with the most frequently reported designations being Assistant Professor (24%) and Consultant Chemical Pathologist (23%). For most respondents (73%), their current employment was directly related to chemical pathology. (Table 1) While 63% reported practising exclusively in chemical pathology, 70% indicated that their professional responsibilities also encompassed other pathology disciplines.

Research and Academic Engagement

Over half (55%) had additional health professions education training. Research involvement was notably high (85%), with all male participants contributing to research compared to 79% of females (p = 0.011). Publication activity peaked during 2024–2025, accounting for 64% of recent outputs. The index category (≥75), reflecting higher research output or quality, was significantly more frequent among participants as determined byFisher's exact test (p=0.018). One-third (32%) were members of international organisations (CAP, IFCC, AACC, etc.), with females being more involved.

Overall Training Satisfaction and Individual Training Components

Overall, respondents rated their FCPS training experience positively, with 30% rating it excellent, 29% as very good, 30% as good and 9% as fair. Most participants rated their academic and professional mentorship provided by the faculty as excellent (40%) or very good (35%) and there was no statistically significant difference across groups. In terms of confidence and leadership development during FCPS training, most participants reported very good (34%) or excellent (39%) improvement. Laboratory skills were rated as very good or excellent by 81% of participants. Ratings for laboratory management exposure were lower, with 69% describing it as very good or excellent. Male respondents were more likely to rate their laboratory management exposure favourably, although this difference did not reach statistical significance (p=0.065). Gender differences were significant in graduation year (p=0.049), work setting (p=0.034), research involvement (p=0.011), and index category (<75 vs ≥75) (p =0.018). For most other factors, such as employment, designation and training quality, no significant gender effect was observed.

Composite Domain Analysis of Training Experience

To provide a more integrated assessment of training experiences, Likert-scale items were grouped into composite domains. The Training Quality domain (laboratory skills, confidence, and leadership development) demonstrated consistently high scores across respondents, with an overall mean score of approximately 4.1 out of 5. The Academic Mentorship and Research domain (faculty mentorship, research opportunities, and professional growth support) also showed high overall ratings (mean ≈ 3.9–4.0 out of 5) as depicted in Table 2 where values are presented as mean ± standard deviation. Differences between genders were not statistically significant.

In contrast, the Laboratory Management Exposure domain consistently scored lower than other domains across all respondent groups, indicating a relative gap in management and administrative training. Gender-based comparisons showed slightly higher mean scores among male respondents across domains. The heat map further illustrated clear patterns across composite domains, with stronger training and mentorship scores among research-active respondents and persistently lower laboratory management exposure across demographic groups (Figure 3).

Graduation Cohort–Based Trends

When respondents were stratified by graduation cohort (pre-2010, 2010–2019, and ≥2020), a trend toward progressively higher training satisfaction and mentorship scores was observed among more recent graduates. Respondents graduating in 2020 or later reported the highest mean scores for both Training Quality and Academic Mentorship domains, while those graduating before 2010 reported comparatively lower scores.

These findings suggest gradual improvement in FCPS training structure and educational support over time, potentially reflecting increased emphasis on structured assessments, mentorship, and competency-based training approaches in recent years as shown in Figure 2.

Association Between Research Involvement and Training Outcomes

Respondents actively engaged in research demonstrated higher perceived training quality and stronger academic mentorship experiences compared to those not involved in research. Research-active participants reported higher mean scores for the Training Quality domain (4.13 vs 3.88) and the Academic Mentorship and Research domain (4.00 vs 3.50).

Table 1: Demographics of FCPS-trained chemical pathologists in Pakistan.

	Total	Female	Male
	N=82	N=56	N=26
Age Group	Age Group	Age Group	Age Group
31-40	42 (51%)	30 (54%)	12 (46%)
41-50	23 (28%)	16 (29%)	7 (27%)
51-60	7 (9%)	6 (11%)	1 (4%)
≥60	8 (10%)	2 (4%)	6 (23%)
Graduation year	Graduation year	Graduation year	Graduation year
1978-1988	6 (8%)	1 (2%)	5 (21%)
1999-2009	25 (35%)	17 (35%)	8 (33%)
2010-2020	40 (56%)	29 (60%)	11 (46%)
≥ 2021	1 (1%)	1 (2%)	0 (0%)
Province of residence			
Balochistan	1 (1%)	0 (0%)	1 (4%)
Gilgit Baltistan	1 (1%)	0 (0%)	1 (4%)
KPK	6 (7%)	3 (5%)	3 (12%)
Outside Pakistan	5 (6%)	5 (9%)	0 (0%)
Pakistan (Province not disclosed)	3 (4%)	2 (4%)	1 (4%)
Punjab	47 (57%)	33 (59%)	14 (54%)
Sindh	19 (23%)	13 (23%)	6 (23%)
Institution where FCPS chemical pathology training was completed?	Institution where FCPS chemical pathology training was completed?	Institution where FCPS chemical pathology training was completed?	Institution where FCPS chemical pathology training was completed?
AFIP, Rawalpindi	26 (32%)	11 (20%)	15 (58%)
Chughtai institute of pathology, Lahore	3 (4%)	2 (4%)	1 (4%)
CMH Lahore	4 (5%)	3 (5%)	1 (4%)
CMH Malir Cantt, Karachi	2 (2%)	2 (4%)	0 (0%)
DIMC, Ojha Campus, Karachi	1 (1%)	0 (0%)	1 (4%)
Dr. Ziauddin medical university & hospital	3 (4%)	2 (4%)	1 (4%)
Indus Hospital, Karachi	1 (1%)	1 (2%)	0 (0%)

	Total	Female	Male
Liaquat national hospital	2 (2%)	2 (4%)	0 (0%)
Nishtar Hospital, Multan	1 (1%)	1 (2%)	0 (0%)
PNS Shifa	2 (2%)	2 (4%)	0 (0%)
Current employment status	Current employment status	Current employment status	Current employment status
Employed full-time	65 (79%)	41 (73%)	24 (92%)
Employed part-time	11 (13%)	9 (16%)	2 (8%)
Homemaker	1 (1%)	1 (2%)	0 (0%)
In Training	1 (1%)	1 (2%)	0 (0%)
Not Employed	2 (2%)	2 (4%)	0 (0%)
Visiting Faculty	1 (1%)	1 (2%)	0 (0%)
Current designation	Current designation	Current designation	Current designation
APWMO	4 (5%)	4 (7%)	0 (0%)
Assistant Professor	20 (24%)	15 (27%)	5 (19%)
Associate Professor	3 (4%)	2 (4%)	1 (4%)
CEO/Principal/Professor	1 (1%)	0 (0%)	1 (4%)
Chemical Pathologist	1 (1%)	1 (2%)	0 (0%)
Consultant	19 (23%)	15 (27%)	4 (15%)
HOD	5 (6%)	1 (2%)	4 (15%)
Pathologist	6 (7%)	5 (9%)	1 (4%)
Type of work setting	Type of work setting	Type of work setting	Type of work setting
Rural	3 (4%)	0 (0%)	3 (12%)
Semi-Urban	6 (7%)	4 (7%)	2 (8%)
Urban	73 (89%)	52 (93%)	21 (81%)

**Percentages are calculated column-wise within gender categories unless otherwise stated. Percentages may not total 100% due to rounding. APWMO: Assistant Professor / Women Medical Officer.*

Figure 1: Comparison between the current employment sector stratified by gender.

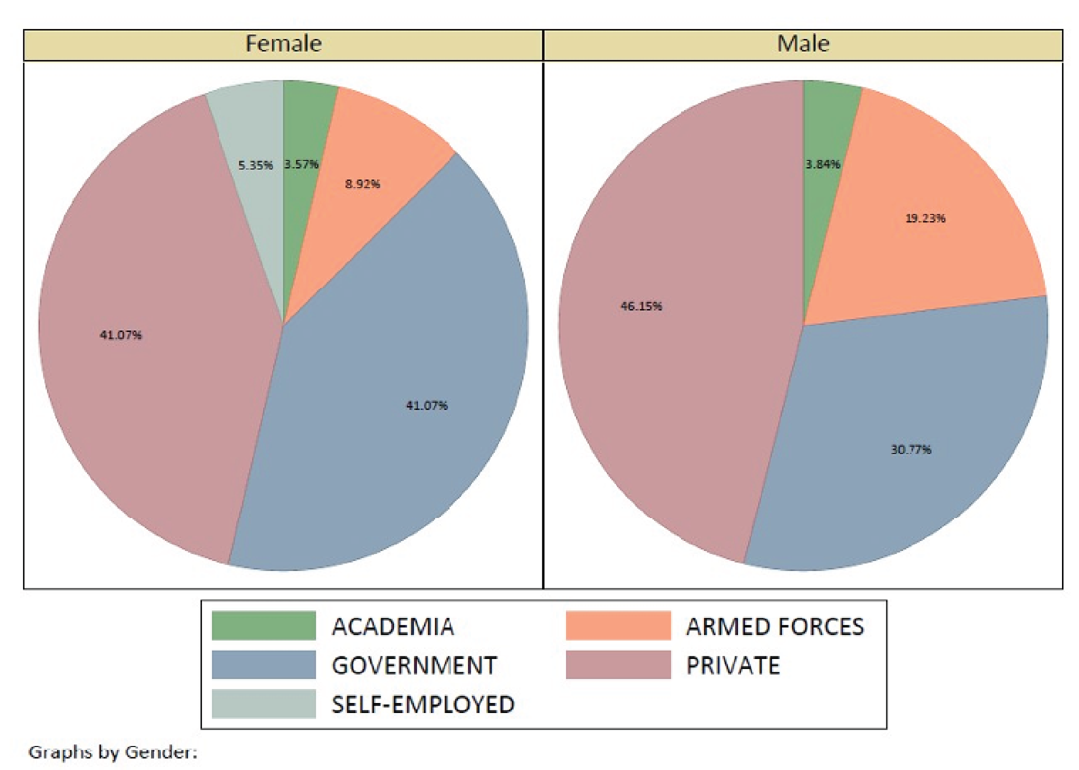


Figure 2: Training quality and academic mentorship score across graduation cohorts.

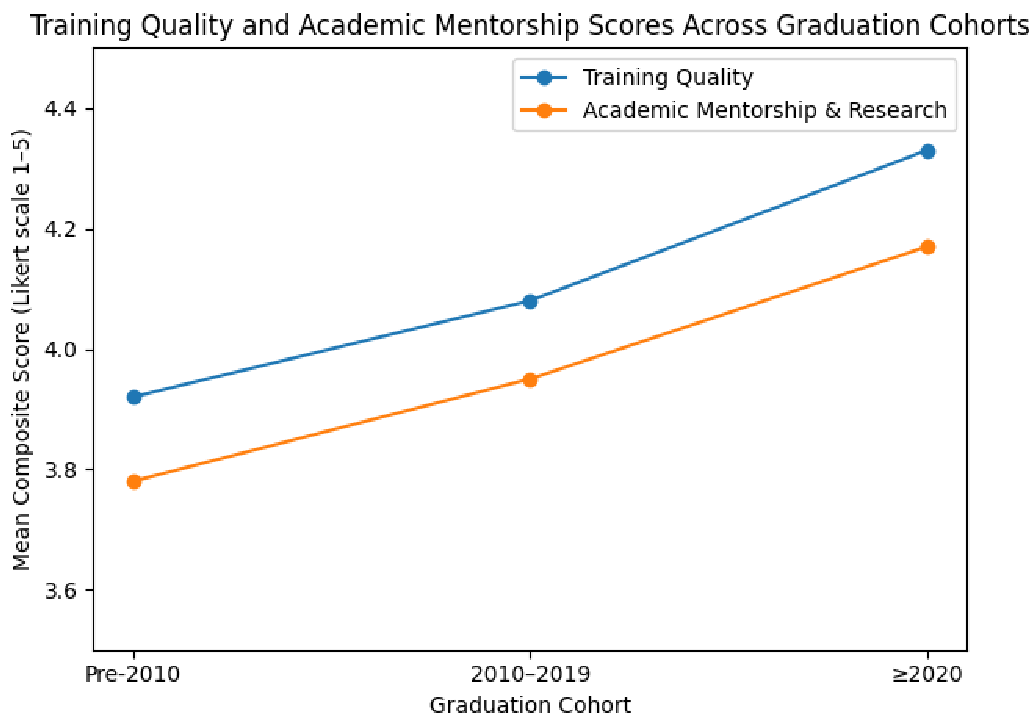


Figure 3: Heat map of composite training domain scores by gender and research involvement.

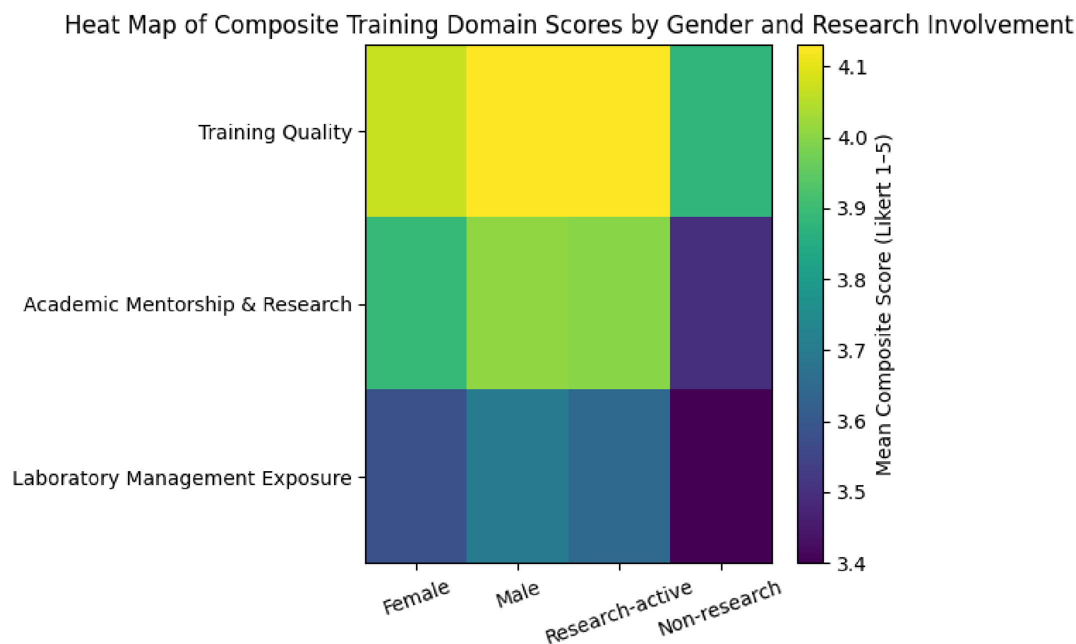


Table 2: Composite Domain Scores of FCPS Training Experience (Likert scale: 1 = very poor, 5 = excellent).

Composite domain	Total (N = 82) Mean ± SD	Female (n = 56) Mean ± SD	Male (n = 26) Mean ± SD
Training Quality*	4.10 ± 0.52	4.07 ± 0.54	4.13 ± 0.49
Academic Mentorship & Research†	3.95 ± 0.60	3.89 ± 0.62	4.01 ± 0.57
Laboratory Management Exposure‡	3.62 ± 0.71	3.58 ± 0.73	3.71 ± 0.63

**Training Quality domain includes laboratory skills, confidence development, and leadership development. †Academic Mentorship & Research domain includes faculty mentorship, research opportunities, and professional growth support. ‡Laboratory Management Exposure domain includes exposure to laboratory administration and management responsibilities.*

Discussion

This cross-sectional study provides an overview of the training and current practices among FCPS-certified chemical pathologists in Pakistan. A higher proportion of responses were obtained from early- and mid-career professionals. This likely reflects greater engagement of younger consultants and trainees with digital communication platforms through which the survey was distributed. Senior consultants, particularly those in administrative or leadership roles, may be less active on social media or electronic survey platforms, leading to under-representation. In addition, the specialty of chemical pathology in Pakistan has expanded primarily over the past three to four decades, and historically only a limited number of training positions were available; consequently, the number of senior practitioners is relatively small compared with the growing cohort of recently trained specialists. Similar response patterns have been reported in other web-based physician surveys. Our study revealed that there is a female predominance, comprising nearly two-thirds of the respondents, which is consistent with the literature that pathology is a female-dominant speciality with thethird-greatest female-to-male proportion of any medical speciality [3]. A study by Samkari et al. reported that female medical students showed more interest in pursuing pathology as compared to their male counterparts [4]. Pathology is well sought after by women due to the reduced number of working hours, fewer on-call responsibility and the flexibility [5]. Most respondents in our study were based in urban centres, consistent with previous reports of uneven distribution of specialists across Pakistan [6]. This disparity stems from the lack of resources, proper living conditions and opportunities and personal career advancement in the rural areas, forcing physicians to work in urban centres [7]. Expanding postgraduate training and employment opportunities in under-represented provinces may help address regional workforce imbalances. Interestingly, our study revealed that females were the only ones working in the self-employed sector. This may reflect the limited availability of flexible institutional positions for women in this speciality in Pakistan, prompting some to pursue independent practice or a preference for self-directed work environments that accommodate personal and family responsibilities. Most of therespondents were employed full-time and pursuing their

careers, highlighting the ever-growing need for chemical pathologists in Pakistan.

Research involvement was notably high (85%), with all male participants contributing to research compared to 79% of females (p = 0.011). This imbalance could be associated with differential access to mentorship and institutional support for female residents [8.9]. Evidence has shown that institutional support and structured research exposure during residency play an important role in empowering and fostering a culture of academic excellence, ultimately helping the residents in career preparation [10]. Encouraging structured research mentorship and equitable access to research infrastructure could therefore help bridge this gender gap. These findings provide evidence to inform CPSP and training institutions on targeted improvements in mentorship and leadership development.

More than half of the participants had additional qualifications in health professions education, suggesting a strong interest in academic and pedagogical development. This trend aligns with the evolving role of chemical pathologists as both educators and consultants, integrating laboratory science with clinical decision-making [11]. The majority of respondents reported high satisfaction with their FCPS training experience and mentorship quality, reflecting the strength of the College of Physicians and Surgeons Pakistan (CPSP) training framework. This is in contradiction to the finding reported by Mahmood et al. that most of the postgraduate trainees were not satisfied with the facilities and mentorship regardless of the training specialities [12]. High ratings in laboratory skills, confidence, and leadership development suggest that current training programs are effectively equipping trainees for independent professional roles. The composite domain analysis provides a more nuanced understanding of FCPS training experiences and highlights important strengths and gaps within current training frameworks. While overall training quality and academic mentorship were rated highly, the consistently lower scores for laboratory management exposure suggest an area requiring targeted enhancement [13]. Similar gaps in management and administrative training have been reported internationally and may limit preparedness for leadership roles in laboratory medicine. The observed improvement in training and mentorship scores among more recent graduates further supports the notion of a progressively evolving FCPS training structure, likely reflecting the increasing incorporation of structured assessments, competency-based education, and formal mentorship models. Collectively, these findings emphasize the need to strengthen laboratory management training and research mentorship within postgraduate chemical pathology programs to ensure comprehensive workforce development.

Although detailed income data were not formally collected in this survey due to sensitivity and anticipated response bias, anecdotal and institutional observations suggest differences between public and private sector employment. In Pakistan, public sector consultants often have structured pay scales with greater job security but limited financial incentives, whereas private sector positions may offer higher compensation but with greater workload variability and performance-based expectations. Thesedifferences may influence career preferences and job satisfaction but require a dedicated workforce economic study to evaluate objectively. Future studies focusing specifically on compensation, job security, and workload distribution across laboratory specialties would be valuable.

This study has several limitations. The participation bias is possible, with respondents more involved in research being more likely to respond. The response rate could not be formally calculated because the survey was disseminated through professional networks and alumni messaging platforms; therefore, the true denominator of eligible participants could not be ascertained, introducing a potential risk of non-response bias. Nevertheless, despite these limitations, our study is one of the only few studies evaluating the current practices and roles in the under-represented, evolving profession of chemical pathology in Pakistan. It underscores the importance of equitable gender participation, structured research mentorship, and continuous professional development in laboratory medicine. Future studies could explore longitudinal outcomes of FCPS trainees, including career trajectories, academic productivity, and international collaborations, to guide evidence-based reforms in postgraduate medical education.

Conclusion

This study highlights the evolving landscape of chemical pathology in Pakistan, characterised by strong female representation, high research engagement, and overall satisfaction with FCPS training. The findings emphasise the need to strengthen institutional support for research activities, especially for women, and promote equitable opportunities for mentorship and leadership, which would further prepare chemical pathologists for leadership roles in clinical diagnostics and academia. Future initiatives should focus on benchmarking local training outcomes against international standards and fostering global collaborations to further strengthen the role of chemical pathologists in clinical care, research, and education.

Ethical Consideration

The survey questionnaire was approved by the Pakistan Society of Chemical Pathologist (PSCP) for ethical considerations. Participation was voluntary, and respondents provided informed consent before completing the survey. No identifying personal data were collected.

Funding

None.

Conflicts of Interests

All authors declare no conflicts of interests.

Author Statements

Sibtain Ahmed: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Supervision.
Amal Mahmood: Literature review, Visualization, Writing – original draft.
Aamir Ijaz: Methodology, Writing – original draft.
Nimarta Heman: Methodology, Writing – original draft.

Ayesha Sadiqa: Formal analysis, Data curation, Writing – review & editing.
Imran Siddiqui: Methodology, Data curation, Writing – review & editing.

Data Availability Statement

Anonymized Data is available on request from the corresponding author.

References

1. Mak CM. Survey on chemical pathology teaching in medical curricula: Insight into patient safety. Biomed J Sci Tech Res [Internet]. 2018;2(3). doi: 10.26717/BJSTR.2018.02.000753
2. Sibtain A, Muhammad Asad KF, Hafsa M, Habib KA, Imran S, Farooq G, et al. From practice to proficiency: Evaluation of a novel workplace-based assessment (WBA) in a postgraduate Chemical Pathology residency program in Pakistan. EJIFCC. 2024;35(2):111–119.doi: 10.56986/ejifcc.v35i2.973
3. Xu H, Remick DG. Pathology: A satisfying medical profession. Acad Pathol. 2016;3:2374289516661559. doi: 10.1177/2374289516661559
4. Samkari A, Qari TA, Abukhairat Jnr SM, AlQarni G, Assiri R, Boraie SA, et al. The attractive and deterring factors for medical students towards pathology as a specialty in Saudi Arabia. Adv Med Educ Pract. 2025;16:1217–1227. doi: 10.2147/AMEP.S482742
5. Tanvir I, Hassan A, Alahmadi S, Waseem H, Anwer J, Shafie A, et al. Ethnic and gender diversity in pathology: A dream deferred. Cureus. 2023;15(5):e38528. doi: 10.7759/cureus.38528
6. Bakhshi SK, Shah Z, Khalil M, Khan Mughal MA, Kazi AM, Virani QUA, et al. Geographical distribution of neurosurgeons andemergency neurosurgical services in Pakistan. World Neurosurg. 2023;179:e515–e522. doi: 10.1016/j.wneu.2023.08.062
7. Sultana A, Awais S, Mughal A, Anwar B. Factors affecting willingness of doctors to work in rural areas of Pakistan. Pak J Public Health. 2017;7(2):100–104. DOI not available
8. Vernuccio F, Arzanauskaite M, Turk S, Torres ET, Choa JMD, Udare AS, et al. Gender discrepancy in research activities during radiology residency. Insights Imaging. 2019;10(1):125. doi: 10.1186/s13244-019-0814-5
9. Yang G, Villalta JD, Weiss DA, Carroll PR, Breyer BN. Gender differences in academic productivity and academic career choice among urology residents. J Urol. 2012;188(4):1286–1290. doi: 10.1016/j.juro.2012.06.023
10. Jeyaraman M, Mariappan T, Jeyaraman N, Nallakumarasamy A, Muthu S. Structured research training in orthopaedics: A systematic review evaluating its role in career development and academic success. J Clin Orthop Trauma. 2025;70:103172. doi: 10.1016/j.jcot.2025.103172
11. Schmidt RL, Garcia CA, Panlener J, Ashwood ER, Jackson BR, Hussong JW. An analysis of clinical consultation activities in clinical chemistry: implications for transformation and resident training in chemical pathology. Arch Pathol Lab Med. 2014;138(5):671–677. doi: 10.5858/arpa.2013-0244-OA
12. Mahmood S, Khaliq T, Raza A. Post graduate training. Pak Armed Forces Med J. 2001;68:807–811. Available from: https://www.pafmj.org/index.php/PAFMJ/article/view/87 (Accessed on: 29/12/2025)
13. Weiss RL, Hassell LA, Parks ER. Progress toward improved leadership and management training in pathology. Archives of Pathology & Laboratory Medicine. 2014;138(4):492-497. doi: https://doi.org/10.5858/arpa.2013-0288-RA

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Assessing Testosterone Epidemiology Using Distributional Approaches: Evidence From Large-Scale Longitudinal Laboratory Dataset in Pakistan

Sibtain Ahmed[1*], Minal Imran[2], Imran Siddiqui[1]

[1]*Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*
[2]*Aga Khan University Medical College, Karachi, Pakistan*

Article Info

[*]***Corresponding Author:***
Sibtain Ahmed
Assistant Professor & Consultant Chemical Pathologist
The Aga Khan University
Stadium Road, P.O. Box 3500, Karachi 74800, Pakistan
Telephone: +92 21 34861951
E-mail: Sibtain.ahmed@aku.edu

Keywords

Testosterone, Trends, Pakistan, Laboratory, Population

Abstract

Background: Testosterone is the main circulating androgen in adult males and serves a crucial role in reproductive function, skeletal health, metabolic regulation, and overall physiological stability. In recent decades, concern has emerged regarding possible secular changes in serum testosterone concentrations, independent of aging. However, findings across populations remain inconsistent, and most evidence originates from high-income settings.

Objective: This study aimed to evaluate long-term temporal patterns in serum testosterone levels among adult men, with particular emphasis on distributional shifts and population-wide hormonal instability, rather than solely on a mean decline.

Methods: A retrospective observational analysis was conducted using anonymized laboratory-based serum testosterone measurements from adult males aged ≥18 years between 2008 and 2025. Only the earliest measurement per individual was included to reduce repeat testing bias. Annual descriptive statistics were calculated, and temporal trends were assessed using linear regression. Age-stratified analysis, percentile drift evaluation, and reference interval shifts were used to explore changes across the testosterone distribution over time.

Results: A total of 104,559 testosterone measurements were analysed over 17 years. Linear regression demonstrated a statistically significant but clinically minimal association between calendar year and testosterone concentration ($\beta = +1.97$ ng/dL/year, $p < 0.001$; R2 = 0.002). Substantial overlap in testosterone distribution was observedacross age groups, with notable inter year variability. Percentile drift analyses indicated parallel movement across the distribution, suggesting population-wide hormonal fluctuation rather than a uniform secular decline.

Conclusion: Mean testosterone levels did not demonstrate a consistent age-independent decline. Instead, findings support a temporal instability and persistent low exposure subgroups, emphasising the importance of distributional approaches in understanding contemporary hormonal epidemiology.

Introduction

In adult males, testosterone serves as the dominant circulating androgen and is essential for normal reproductive physiology as well as for maintaining skeletal strength, muscle integrity, and metabolic function [1]. Beyond its reproductive effects, testosterone contributes to bone mineralization, preservation of lean body mass, regulation of erythropoiesis, and cardiometabolic homeostasis throughout adulthood [2, 3]. Reduced testosterone concentrations have been associated with adverse outcomes such as osteoporosis, metabolic dysfunction, frailty, and increased overall morbidity, prompting greater clinical and public health interest in long-term hormonal patterns among men [4].

Over recent decades, concern has emerged regarding potential age-independent changes in serum testosterone concentrations in male populations[5,6]. Multiple hypotheses have been proposed to explain these observations, including increasing rates of obesity, exposure to environmental endocrine-disrupting chemicals, reduced physical activity, psychosocial stressors, and evolving utilization patterns [7]. However, published findings have varied substantially, with no consistent evidence supporting a uniform or progressive decline across all populations. A major limitation of prior research is its reliance on population averages, which may fail to capture clinically important changes occurring at different points within the testosterone distribution, particularly among individuals with persistently low levels. In addition, most available data are derived from high-income countries, limiting the applicability of existing findings to other regions with distinct demographic and environmental characteristics [8].

To address these limitations, the present study analyses serum testosterone measurements obtained from a large laboratory-based cohort over a 17-year period. By evaluating age-stratified trends, distributional shifts, percentile drift, and temporal changes in reference intervals, this study seeks to determine whether observed hormonal patterns reflect a simple linear decline or a broader population-level instability with earlier exposure to suboptimal testosterone levels.

Methods

This study was a retrospective observational study based on serum testosterone measurements obtained from a large clinical laboratory database. The dataset included all adult male testosterone tests performed between 2008 and 2025. All data were anonymized prior to analysis, and no patient level identifiers were accessible. The Aga Khan University (AKU) Laboratory receives specimens through a nationwide network of 13 outreach laboratories and more than 300 collection centers across Pakistan, encompassing urban and rural regions. This extensive coverage supports the representativeness of the study population with respect to the broader adult male population of Pakistan. The study was approved by the institutional ethical review committee of Aga Khan University (2025-11907-36009).

The study population consisted of adult men aged 18 years and above who underwent serum testosterone testing during the study period. Where multiple measurements were available for the same individual, only the earliest recorded test was used for analysis in order to minimize bias related to repeat testing andclinical follow-up. The age at the time of testing and the calendar year of measurement were recorded for all included samples. As this was a laboratory-based retrospective dataset, detailed clinical variables such as thyroid status or other endocrine diagnoses were not consistently available. Therefore, no additional filtering based on hormonal comorbidities was performed, and the analysis was intended to capture population-level distributional patterns of testosterone within routine clinical testing Serum testosterone concentrations were measured as part of routine clinical care using standardized laboratory assays. Serum total testosterone concentrations were measured using electrochemiluminescence immunoassay (ECLIA) on Roche diagnostic platforms during the study period. Testing was initially performed on the Roche Elecsys e170 analyzer, subsequently on the Elecsys e411 platform, and currently on the Cobas e801 analyzer. Results were reported in nanograms per deciliter. Although assay platforms evolved over the study period, all measurements were performed in accredited laboratory settings with established internal and external quality control procedures. Descriptive statistics were calculated annually, including sample size, mean, median, and standard deviation. Temporal trends were assessed using linear regression, with testosterone concentration as the dependent variable and calendar year as the independent variable. Effect size, statistical significance, and explained variance (R^2) were reported to distinguish statistical from clinical relevance. Age-stratified analyses were performed to minimize confounding by aging. Distributional changes were assessed using percentile drift analyses, comparisons between early and late testing cohorts, and longitudinal evaluation of reference interval shifts. Analyses were conducted using Python (pandas, matplotlib). A two-sided p-value <0.05 was considered statistically significant.

Results

A total of 104,559 adult male testosterone measurements were analysed over a 17-year period (2009–2025). Linear regression demonstrated a statistically significant but clinically minimal association between calendar year and testosterone concentration ($\beta = +1.97$ ng/dL/year, $p < 0.001$; $R^2 = 0.002$), indicating that calendar time alone explained a negligible proportion of hormonal variability.

Age-stratified analyses revealed substantial overlap in testosterone distributions between younger and older men, with repeated periods in which younger adults exhibited testosterone levels comparable to those traditionally associated with ageing. Temporal fluctuations were observed across all age groups, suggesting population-wide instability rather than age-restricted decline.

These findings indicate that loss of hormonal robustness and earlier exposure to suboptimal testosterone levels, rather than simple linear decline, may represent the dominant contemporary pattern. Percentile drift analyses revealed parallel movement across the entire testosterone distribution, suggesting population-wide hormonal changes. Early versus late cohort comparisons demonstrated shifts in distributional profiles, while reference interval analyses showed temporal variability in both lower and upper limits as shown in Table

1. Overlayedhistograms comparing testosterone distributions between early and late testing cohorts, supporting temporal shifts beyond age-related effects as shown in Figure 1. Longitudinal changes in the 10th, 25th, 50th, 75th, and 90th percentiles of serum testosterone across calendar years, demonstrating population-wide distributional drift as depicted in Figure 2. Year-wise movement of the lower (2.5th percentile) and upper (97.5th percentile) reference limits, highlighting instability of static reference intervals over time as described in Figure 3. Prevalence of Low Testosterone (<300 ng/dL) by year is defined in Table 2.

Table 1: Annual Distribution of Adult Male Testosterone Levels.

Year	n	Mean (ng/dL)	Median (ng/dL)	SD
2008	4352	369.9	348.0	204.8
2009	4739	389.2	363.4	221.3
2010	5128	377.4	349.2	218.6
2011	5031	378.2	354.5	215.1
2012	5012	382.5	356.5	210.3
2013	5045	365.1	333.2	210.6
2014	5046	405.2	377.4	218.7
2015	4915	387.0	361.8	216.8
2016	5600	386.5	355.5	223.1
2017	5797	397.4	367.8	218.7
2018	6035	390.2	364.0	216.2
2019	6522	407.3	381.5	219.1
2020	5698	402.8	379.1	208.7
2021	6786	401.1	374.3	209.8
2022	6467	390.2	366.4	203.7
2023	7310	393.2	370.0	202.5
2024	8178	406.0	380.0	207.8
2025	6898	421.0	398.0	211.2

Table 2: Prevalence of Low Testosterone (<300 ng/dL) by Year.

Year	Total tests (n)	<300 ng/dL (n)	<300 ng/dL (%)
2008	4352	1673	38.4
2009	4739	1678	35.4
2010	5128	1974	38.5
2011	5031	1857	36.9
2012	5012	1802	36.0
2013	5045	2093	41.5
2014	5046	1634	32.4
2015	4915	1757	35.7
2016	5600	2070	37.0
2017	5797	2008	34.6

Year	Total tests (n)	<300 ng/dL (n)	<300 ng/dL (%)
2018	6035	2139	35.4
2019	6522	2109	32.3
2020	5698	1773	31.1
2021	6786	2182	32.2
2022	6467	2228	34.5
2023	7310	2390	32.7
2024	8178	2495	30.5
2025	6898	1914	27.7

Figure 1: Early vs Late Cohort Testosterone Distribution.

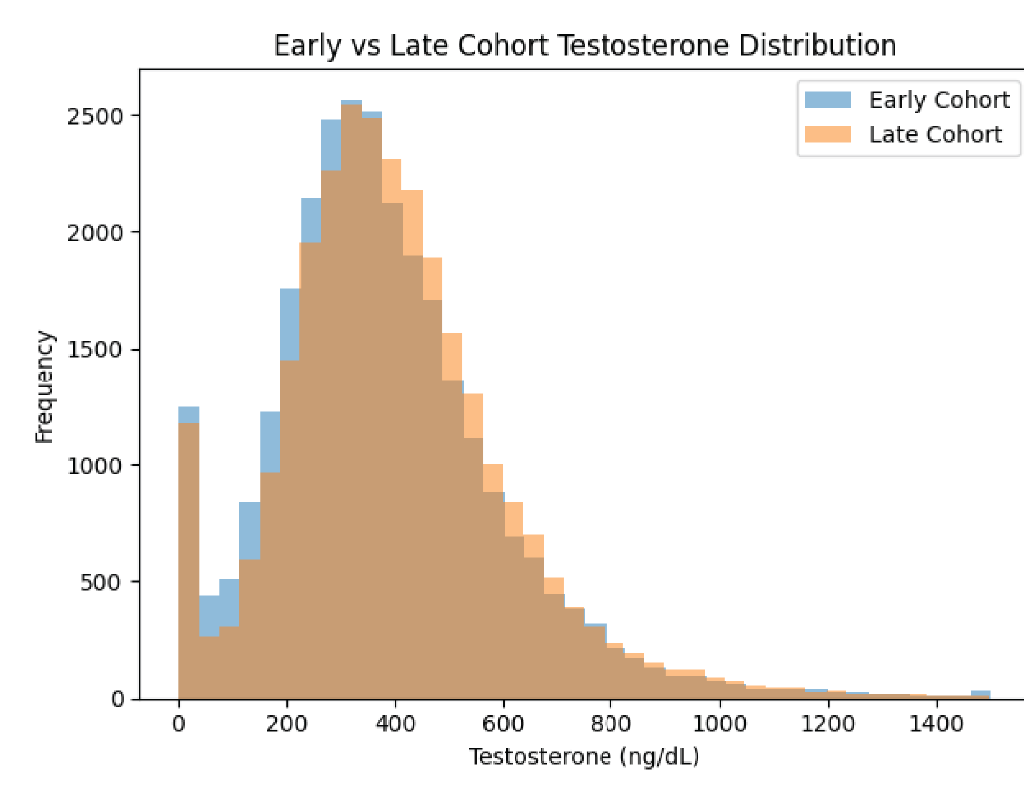


Figure 2: Percentile Drift of Serum Testosterone Over Time.

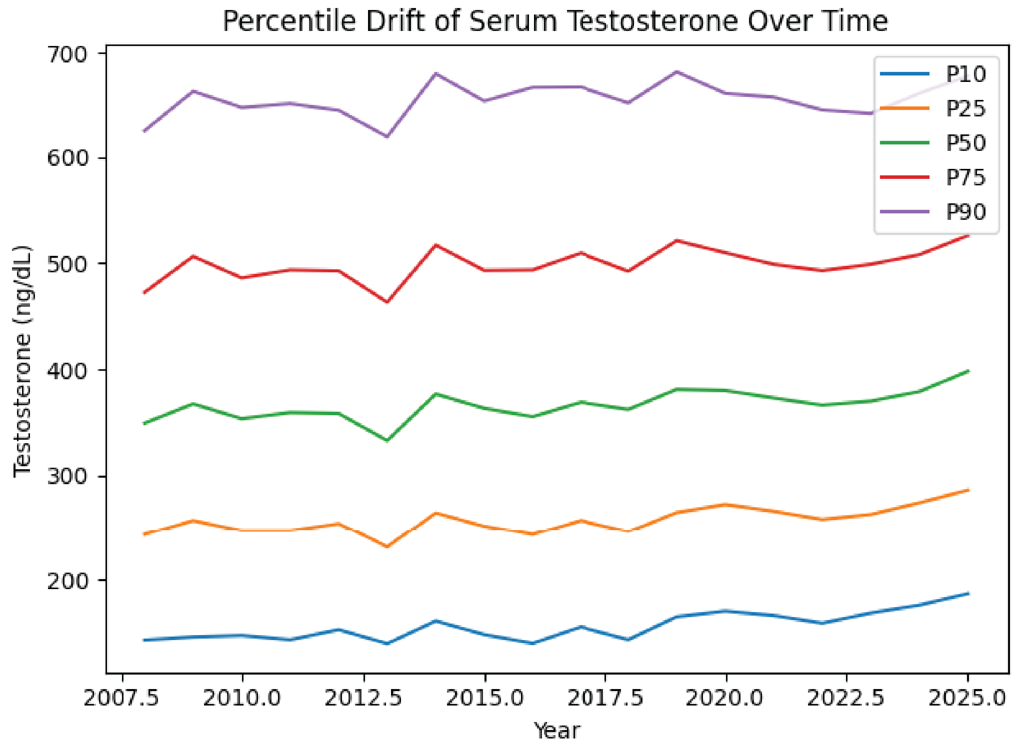
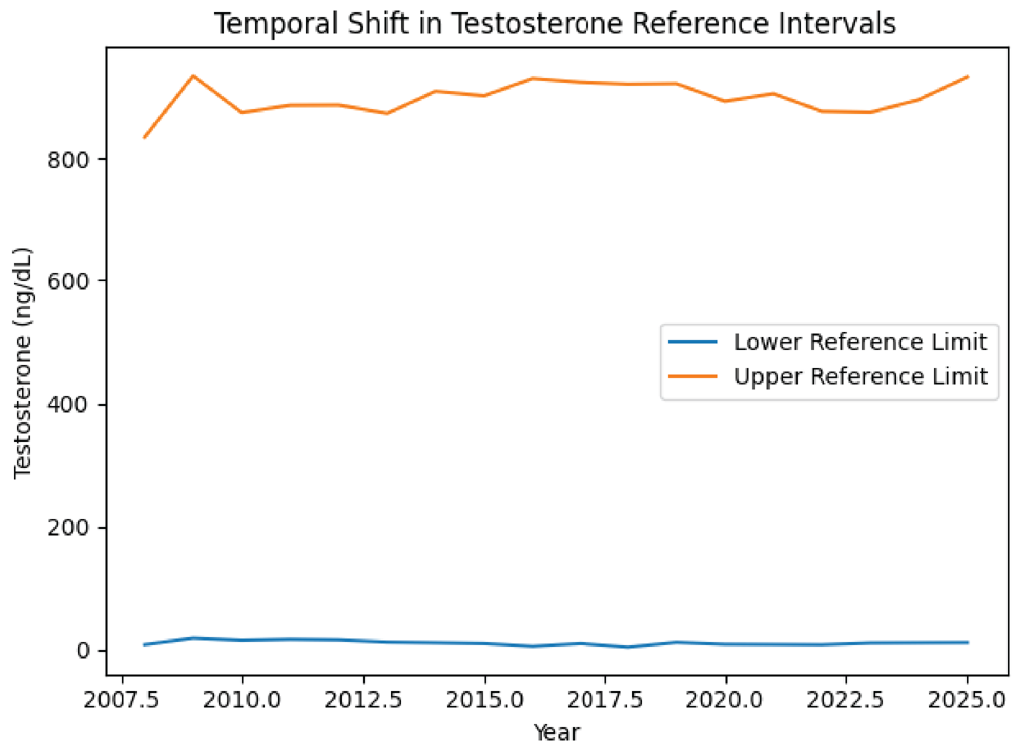


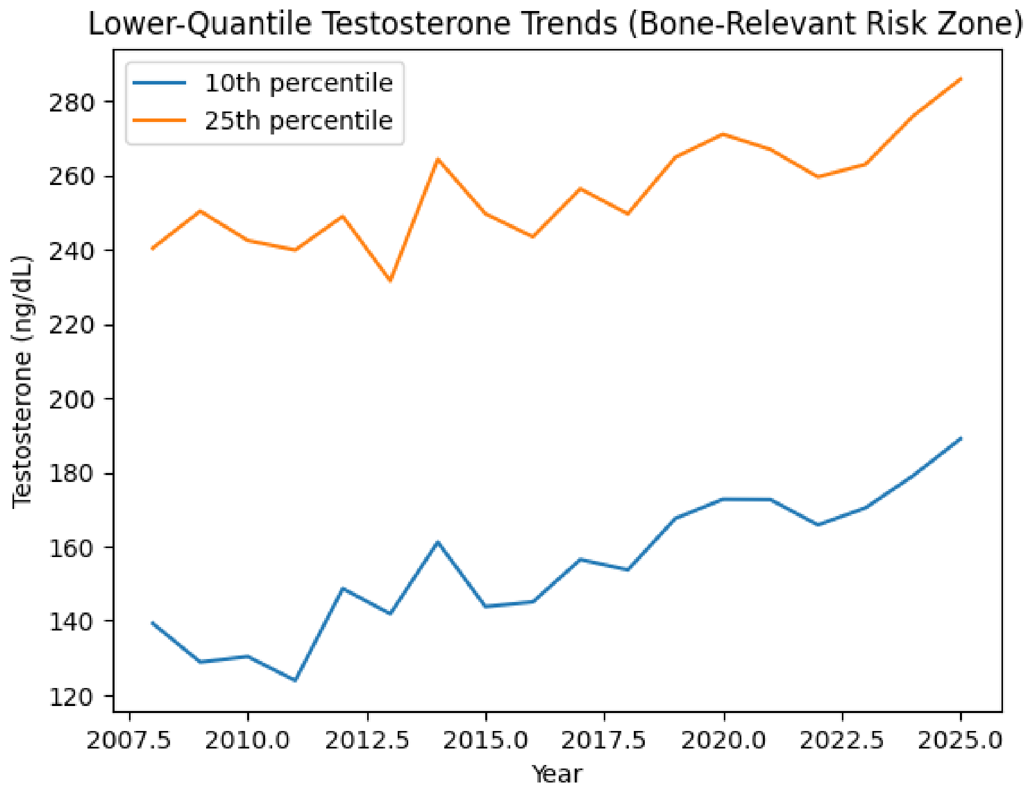
Figure 3: Temporal Shift in Testosterone Reference Intervals.



Wise movement of the lower (2.5th percentile) and upper (97.5th percentile) reference limits, highlighting instability of static reference intervals over time.

The Figure 4 demonstrates marked inter-year variability rather than a monotonic decline, with periods of hormonal suppression followed by partial recovery. Despite fluctuations in mean values,wide dispersion persists across all years, indicating substantial population heterogeneity relevant to long-term skeletal health.

Figure 4: Temporal Trend of Mean Serum Testosterone Levels in Adult Men (2008–2025).



Year-wise trends in the 10th and 25th percentiles of serum testosterone concentrations in adult men. These lower quantiles represent individuals at highest biological risk for impaired bone metabolism and fragility fractures as shown in Figure 5. Persistently low values across the study period highlight a substantial subgroup of men chronically exposed to bone-adverse testosterone levels, even when population mean values appear stable.

Figure 5: Lower-Quantile (10th and 25th Percentile) Testosterone Trends Over Time.

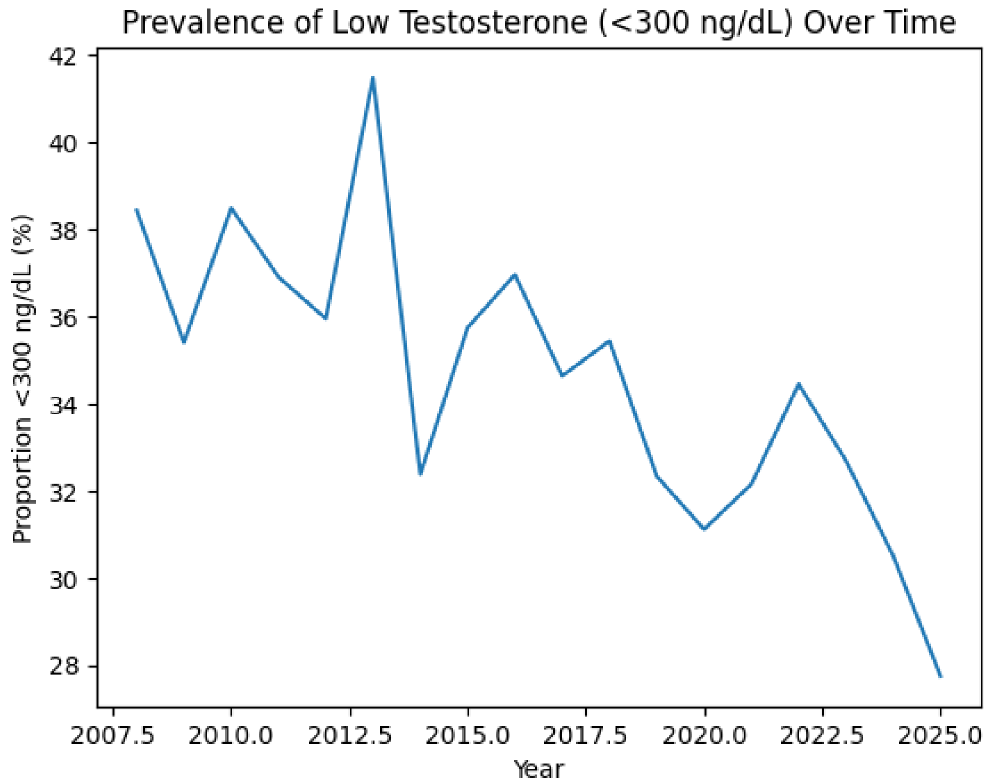
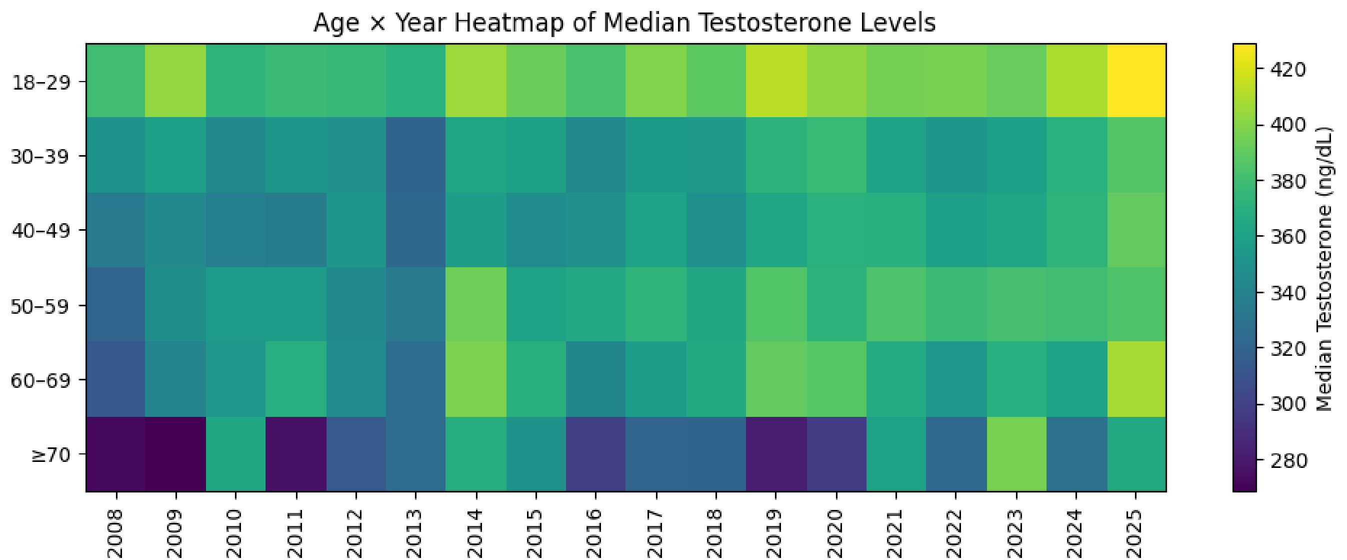


Figure 6: Age × Year Heatmap of Median Serum Testosterone Levels.



Heatmap illustrating median serum testosterone concentrations (ng/dL) across age bands and calendar years.

Discussion

In this large laboratory-based analysis spanning 17 years and over 104,000 adult male testosterone measurements, we found no evidence of a consistent age-independent linear decline in mean serum testosterone. Although linear regression revealed a statistically significant association between calendar year and testosterone, the effect size was small and explained a tiny portion of the overall variability. Instead, variability occurredacross all age groups, with substantial temporal fluctuation and overlapping testosterone distributions, suggesting population-wide instability rather than a uniform decline. Our findings are consistent with some recent literature that emphasizes heterogeneity in testosterone levels. For instance, meta-analysis of testosterone and health outcomes has shown considerable variability in testosterone levels across populations and limited associations with age or outcome endpoints,highlighting the complexity of population-level

hormone dynamics [8]. Such variability supports the notion that mean-level trends may not capture shifts in hormonal distributions [9].

However, some systematic reviews have reported evidence of declining trends in testosterone concentrations over several decades [5,6]. A recent comprehensive literature synthesis identified a negative linear association between testosterone and calendar year across large international cohorts, independent of age and body mass index, suggesting a possible secular decline [10]. These differences may arise from variation in study populations, measurement methods, and analytic frameworks, reinforcing the need for context-specific interpretations.

The present study builds on prior research by incorporating distributional analyses, including percentile drift and reference interval shifts. This approach revealed broad movement across the entire testosterone distribution, meaning changes were not confined to the central tendency alone.[8] Similar distributional concerns have been raised in population hormone studies, where reliance on mean values might obscure important subgroups at risk, such as those with chronically low testosterone [10,11].

Clinically, the presence of a persistent subgroup with low testosterone has important implications [12]. Low testosterone has been linked to adverse outcomes, including impaired bone health, metabolic dysfunction, and frailty [13,14]. Our findings of instability and early exposure to lower serum concentration suggest that timing and distributional context, rather than mean levels alone, may better capture biologically relevant population patterns.

This study has several limitations. First, as a laboratory-based analysis, the sample may reflect healthcare-seeking populations rather than community-based cohorts, which can introduce selection bias. Second, assay changes over time, and a lack of standardized mass spectrometry measurements across the entire period may contribute to analytical heterogeneity [15,16]. Finally, while distributional methods provide nuanced insights, they may be less comparable to traditional epidemiological metrics in other published studies. Moreover, detailed clinical variables such as comorbidities, thyroid status, body mass index, medication use, and lifestyle factors were not consistently available in the laboratory information system and therefore could not be controlled for the analysis.

In conclusion, while simple linear regression suggests only minimal temporal change in testosterone levels, our population-level distributional analyses highlight instability and persistent low exposures across age groups. These findings challenge the interpretation of secular hormonal decline based solely on mean values and emphasize the need for broader distributional perspectives in hormonal epidemiology.

Data Availability

Data is available on reasonable requests from the corresponding author.

Conflicts of interest

None.

Funding

None.

Ethical Approval

The study was approved by the institutional ethical review committee of Aga Khan University (2025-11907-36009).

References

1. Finkelstein JS, Lee H, Burnett-Bowie S-AM, et al. Gonadal steroids and body composition, strength, and sexual function in men. N Engl J Med 2013;369:1011–1022. https://doi.org/10.1056/NEJMoa1206168.
2. Basaria S. Male hypogonadism. Lancet 2014;383:1250–1263. https://doi.org/10.1016/S0140-6736(13)61126-5.
3. Bhasin S, Brito JP, Cunningham GR, et al. Testosterone Therapy in Men With Hypogonadism: An Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab 2018;103:1715–1744. https://doi.org/10.1210/jc.2018-00229.
4. Cruickshank M, Hudson J, Hernández R, et al. The effects and safety of testosterone replacement therapy for men with hypogonadism: the TestES evidence synthesis and economic evaluation. Health Technol Assess 2024;28:1–210. https://doi.org/10.3310/JRYT3981.
5. Fantus RJ, Greenberg DR, Chang C, et al. Novel reference range values for serum testosterone: a cross-sectional study of 200,000 males. J Endocrinol Invest 2024;47:2261–2268. https://doi.org/10.1007/s40618-024-02319-0.
6. Feldman HA, Longcope C, Derby CA, et al. Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts male aging study. J Clin Endocrinol Metab 2002;87:589–598. https://doi.org/10.1210/jcem.87.2.8201.
7. Kim SH, Park JJ, Kim KH, et al. Efficacy of testosterone replacement therapy for treating metabolic disturbances in late-onset hypogonadism: a systematic review and meta-analysis. Int Urol Nephrol 2021;53:1733–1746. https://doi.org/10.1007/s11255-021-02876-w.
8. Mazur A, Westerman R, Mueller U. Is rising obesity causing a secular (age-independent) decline in testosterone among American men? PLoS One 2013;8:e76178. https://doi.org/10.1371/journal.pone.0076178.
9. Krishnan S, Aldana-Bitar J, Golub I, et al. Testosterone therapy and the risk of cardiovascular disease in older, hypogonadal men. Prog Cardiovasc Dis 2024;84:14–18. https://doi.org/10.1016/j.pcad.2024.02.015.
10. Midzak AS, Chen H, Papadopoulos V, et al. Leydig cell aging and the mechanisms of reduced testosterone synthesis. Mol Cell Endocrinol 2009;299:23–31. https://doi.org/10.1016/j.mce.2008.07.016.
11. Pantalone KM, George J, Ji X, et al. Testosterone replacement therapy and the risk of adverse cardiovascular outcomes and mortality. Basic Clin Androl 2019;29:5. https://doi.org/10.1186/s12610-019-0085-7.
12. Sood A, Hosseinpour A, Sood A, et al. Cardiovascular Outcomes of Hypogonadal Men Receiving Testosterone Replacement Therapy: A Meta-analysis of Randomized Controlled Trials. Endocr Pract 2024;30:2–10. https://doi.org/10.1016/j.eprac.2023.09.012.

13. Travison TG, Araujo AB, O'Donnell AB, et al. A population-level decline in serum testosterone levels in American men. J Clin Endocrinol Metab 2007;92:196–202. https://doi.org/10.1210/jc.2006-1375
14. Travison TG, Vesper HW, Orwoll E, et al. Harmonized Reference Ranges for Circulating Testosterone Levels in Men of Four Cohort Studies in the United States and Europe. J Clin Endocrinol Metab 2017;102:1161–1173.
15. Wu FCW, Tajar A, Beynon JM, et al. Identification of late-onset hypogonadism in middle-aged and elderly men. N Engl J Med 2010;363:123–135. https://doi.org/10.1056/NEJMoa0911101.
16. Yeap BB, Marriott RJ, Dwivedi G, et al. Associations of Testosterone and Related Hormones With All-Cause and Cardiovascular Mortality and Incident Cardiovascular Disease in Men : Individual Participant Data Meta-analyses. Ann Intern Med 2024;177:768–781. https://doi.org/10.7326/M23-2781.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Comparative performance analysis of three point-of-care glucose testing devices against a central laboratory reference method

Auliya Ashar[1], Fauqa Arinil Aulia[2,3], Ferdy Royland Marpaung[2,3], M Robiul Fuadi[2,4], Yessy Puspitasari[2,3,4*]

[1]*Clinical Pathology Study Program, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia*
[2]*Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia*
[3]*Department of Clinical Pathology, Dr. Soetomo General Academic Hospital, Surabaya, Indonesia*
[4]*Universitas Airlangga Hospital, Surabaya, Indonesia*

Article Info

[*]***Corresponding Author:***

Yessy Puspitasari
Department of Clinical Pathology, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia
Department of Clinical Pathology, Dr. Soetomo General Academic Hospital, Surabaya, Indonesia
Universitas Airlangga Hospital, Surabaya, Indonesia
E-mail : yessy.puspitasari@fk.unair.ac.id

Keywords

Point-of-care testing, Glucose monitoring, Analytical accuracy, Hexokinase method, ISO 15197

Abstract

Background: Point-of-care testing (POCT) glucose systems are widely used in clinical practice due to their rapid turnaround time and operational convenience. However, concerns remain regarding their analytical agreement with laboratory reference methods, particularly in relation to accuracy and potential clinical impact.

Methods: This cross-sectional analytical study included 196 adult outpatients at Dr. Soetomo General Academic Hospital, Surabaya, Indonesia. Capillary blood glucose was measured using three POCT devices (StatStrip, Accu-Chek Inform II, and Rightest), while venous plasma glucose was analyzed using a hexokinase-based reference method (Alinity C). Analytical performance was evaluated using correlation analysis, Bland–Altman plots, Passing–Bablok regression, and ISO 15197 accuracy criteria.

Results: All POCT devices demonstrated strong correlations with the reference method ($r = 0.967–0.982$; $p < 0.001$). StatStrip showed a small negative bias (−2.89 mg/dL; $p = 0.003$), while Accu-Chek Inform II exhibited a larger deviation (−7.80 mg/dL; $p < 0.001$). In contrast, Rightest showed no statistically significant bias (−0.14 mg/dL; $p = 0.870$) and demonstrated the closest agreement with the reference method. Agreement analysis confirmed minimal systematic and proportional bias for Rightest, whereas Accu-Chek Inform II showed wider variability. ISO 15197 evaluation indicated the highest accuracy for Rightest and lower compliance for Accu-Chek Inform II, particularly at glucose levels <200 mg/dL.

Conclusion: Although all POCT systems showed strong correlations with the reference method, clinically relevant differences in analytical agreement and accuracy were identified. Rightest demonstrated the best overall performance. These findings highlight the importance of evaluating agreement beyond correlation and support the use of POCT primarily for glucose monitoring rather than definitive diagnosis.

Background

The American Diabetes Association (ADA) defines diagnostic and prediabetic criteria based on blood glucose levels in conjunction with the presence or absence of clinical symptoms [1,2]. Glycemic assessment can be performed either in a central laboratory or through point-of-care testing (POCT), which refers to diagnostic methods designed for use in close proximity to the patient [1,3,4]. POCT platforms offer several advantages over conventional central laboratory testing, including faster turnaround times and reduced specimen volume requirements [5]. Point-of-care glucose testing is a diagnostic technology that enables rapid, automated, efficient, and cost-effective detection of blood glucose levels [6–9]. These devices are widely utilized across various healthcare settings, including hospitals, emergency departments, outpatient clinics, intensive care units (ICUs), and ambulatory services [6,7]. Continuous improvements in analytical performance have been reported; however, POCT glucose meters have not yet gained universal acceptance for diagnostic use [10,11].

POCT glucose meters are used in both inpatient and outpatient settings, although their clinical applications may differ. In outpatient or ambulatory care, glucometers are commonly used for routine glucose monitoring and diabetes self-management. In hospital settings, particularly in emergency departments and intensive care units, POCT systems facilitate rapid clinical decision-making for glycemic control. However, critically ill patients often present physiological conditions such as anemia, hypotension, or drug interference that may affect POCT accuracy. Therefore, many clinical guidelines recommend that POCT glucose measurements obtained in hospitalized patients be interpreted cautiously and verified using laboratory methods when clinically indicated.

POCT systems typically analyze whole blood samples, whereas reference laboratory methods measure glucose levels in venous plasma. Physiologically, glucose concentration in venous plasma is approximately 10%–15% higher than that in whole blood, due to differences in water content between plasma and red blood cells. Most POCT devices convert whole blood measurements to plasma-equivalent glucose concentrations using correction factors that account for hematocrit levels in order to improve comparability with laboratory methods [12]. Variability in glucose measurements may also depend on the type of blood sample used, such as capillary versus venous blood [13]. According to the Clinical Laboratory Improvement Amendments (CLIA), POCT glucose results should fall within ±10% or 5.4 mg/dL of those obtained using reference laboratory methods [14]. This study aimed to evaluate the analytical performance of three POCT glucose systems StatStrip (Nova Biomedical), Accu-Chek Inform II (Roche Diagnostics), and Rightest (Bionime) by comparing their results with those obtained using the laboratory hexokinase reference method on the Alinity analyzer (Abbott Laboratories).

Objective

To evaluate the validity of blood glucose point-of-care testing (POCT) devices StatStrip, Accu-Chek Inform II, and Rightest by comparing their performance to the reference hexokinase methodusing the Alinity analyzer as the standard for glucose measurement.

Methods

Study Design and Setting

This analytical observational study employed a cross-sectional design and was conducted between August and October 2023 at the Clinical Pathology Laboratory of Dr. Soetomo General Academic Hospital, Surabaya, Indonesia. Data were collected using simple random sampling. The study population consisted of adult outpatients in order to minimize potential physiological interferences commonly observed in critically ill hospitalized patients.

Three point-of-care testing (POCT) glucose systems StatStrip® (Nova Biomedical), Accu-Chek Inform II® (Roche Diagnostics), and Rightest® (Bionime) were evaluated. These devices were selected because they represent commonly used glucose monitoring systems in routine clinical practice, including hospital and outpatient settings in Indonesia, and employ different enzymatic detection principles. StatStrip® utilizes a glucose oxidase–based electrochemical detection system, Accu-Chek Inform II® uses a mutant quinoprotein glucose dehydrogenase (Mut. Q-GDH) enzyme system, and Rightest® operates using a glucose dehydrogenase (GDH) detection method. The reference laboratory measurement was performed using the Alinity C analyzer (Abbott Laboratories), which applies the hexokinase enzymatic method. Ethical approval for this study was obtained from the Ethics Committee of Dr. Soetomo General Academic Hospital. All POCT measurements were performed according to the manufacturers' instructions.

Only outpatient participants were included in this study to minimize potential confounding factors commonly present in critically ill patients, such as severe hematocrit abnormalities, hypoperfusion, vasopressor therapy, or interfering substances. These conditions are known to affect the analytical performance of POCT glucose meters. By focusing on outpatient individuals with relatively stable physiological conditions, the study aimed to evaluate the analytical agreement between POCT devices and the laboratory reference method under controlled clinical circumstances.

Sample Collection and Procedure

The study population comprised outpatient individuals undergoing routine blood sampling at the Sampling Unit, 1st Floor, Dr. Soetomo General Hospital, Surabaya. Inclusion criteria were patients aged >18 years who provided written informed consent prior to participation. Capillary blood samples were obtained via single-finger puncture and analyzed immediately using the three POCT glucose meters: StatStrip (Nova), Accu-Chek Inform II (Roche), and Rightest (Bionime). Concurrently, venous blood samples were collected using vacutainer tubes containing sodium fluoride (NaF, gray-top tubes) to inhibit glycolysis. Samples were stored in a small manual cooler box containing a single frozen ice gel pack, which maintained an internal temperature of approximately 6°C to 10°C during transport and prior to analysis, in order to preserve glucose stability. Plasma glucose analysis was conducted within twohours

of sample collection. Prior to testing, venous blood samples were centrifuged to separate the plasma. The resulting plasma was then analyzed for glucose concentration using the Alinity C analyzer (Abbott), which served as the reference standard.

Results

This study included 196 participants. The majority were female (110, 56.1%), while males accounted for 86 participants (43.9%). Based on blood glucose levels, 49 participants (25%) had levels <100 mg/dL, 98 (50%) between 100–200 mg/dL, and 49 (25%) >200 mg/dL. The age of participants ranged from 18 to 83 years,with a median of 54 years and a mean of 51.09 ± 12.823 years. Baseline characteristics are summarized in Table 1. Clinically relevant differences were observed between POCT devices and the reference laboratory method (Table 2). StatStrip demonstrated a small but statistically significant negative bias (mean difference −2.89 mg/dL; 95% CI: −4.81 to −0.97; p = 0.003). Accu-Chek Inform II showed a larger negative bias (−7.80 mg/dL; 95% CI: −9.90 to −5.69; p < 0.001), indicating greater deviation from the reference method. In contrast, the Rightest device showed no statistically significant bias (−0.14 mg/dL; 95% CI: −1.80 to 1.52; p = 0.870).

Table 1: Demographic Characteristics and Blood Glucose Categories of Respondents.

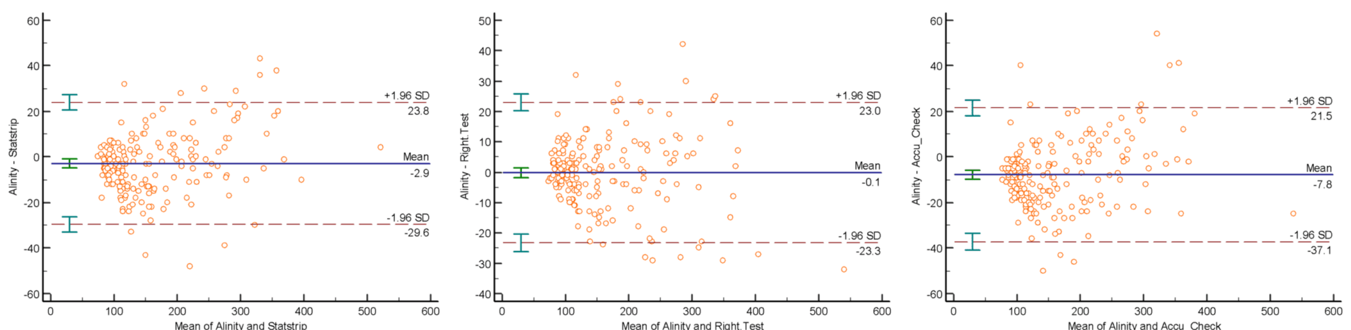
Category	Frequency	Percentage (%)
Male	86	43.9
Female	110	56.1
Total	196	100
< 100 mg/dL	49	25
100-200 mg/dL	98	50
> 200 mg/dL	49	25
Total	196	100
Age (Median [Range])	54 (18 - 83)	Mean ± SD: 51.09 ± 12.823

Table 2: Comparison of Blood Glucose Differences Between StatStrip, Accu-Chek Inform II, and Rightest Against Alinity.

Variable	Mean Difference	95% CI of Difference	P (H: Mean = 0)
Alinity vs StatStrip	-2.89	-4.8131 to -0.9727	0.0033
Alinity vs Accu-Chek	-7.80	-9.9015 to -5.6903	0.0001
Alinity vs Rightest	-0.14	-1.7993 to 1.5238	0.8703

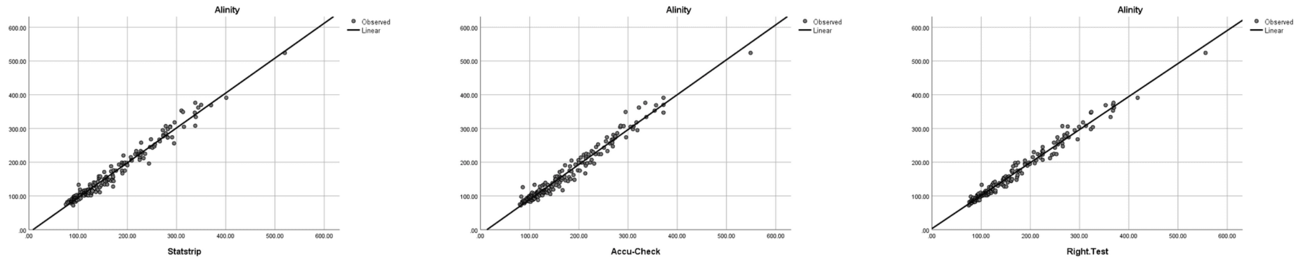
Agreement analysis using Bland–Altman plots demonstrated systematic bias across devices (Figure 1). The Rightest device showed the narrowest limits of agreement, whereas Accu-ChekInform II exhibited wider variability, indicating less consistent agreement with the reference method.

Figure 1: Bland - Altman plots comparing glucose measurements between Alinity and StatStrip A, Rightest B, and Accu-Chek C.



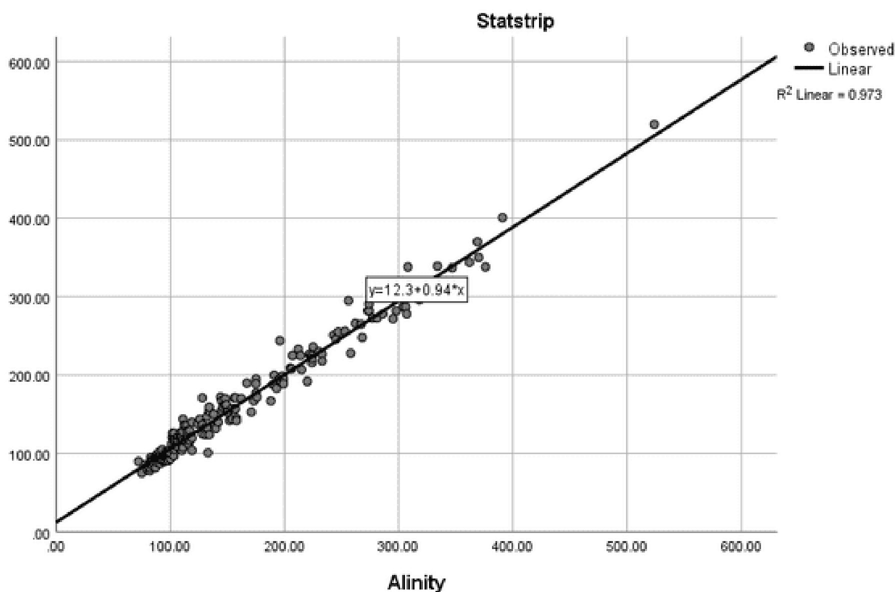
Correlation analysis revealed strong and statistically significant associations between all POCT devices and the reference method (all $p < 0.001$). The highest correlation was observed for the Rightest device ($r = 0.982$), followed by StatStrip ($r = 0.975$) andAccu-Chek Inform II ($r = 0.967$). Scatter plots are presented in Figure 2. However, correlation alone does not imply agreement; therefore, agreement analyses were further performed.

Figure 2: Spearman correlation plots comparing glucose measurements between Alinity and StatStrip A, Accu-Chek B, and Rightest C.



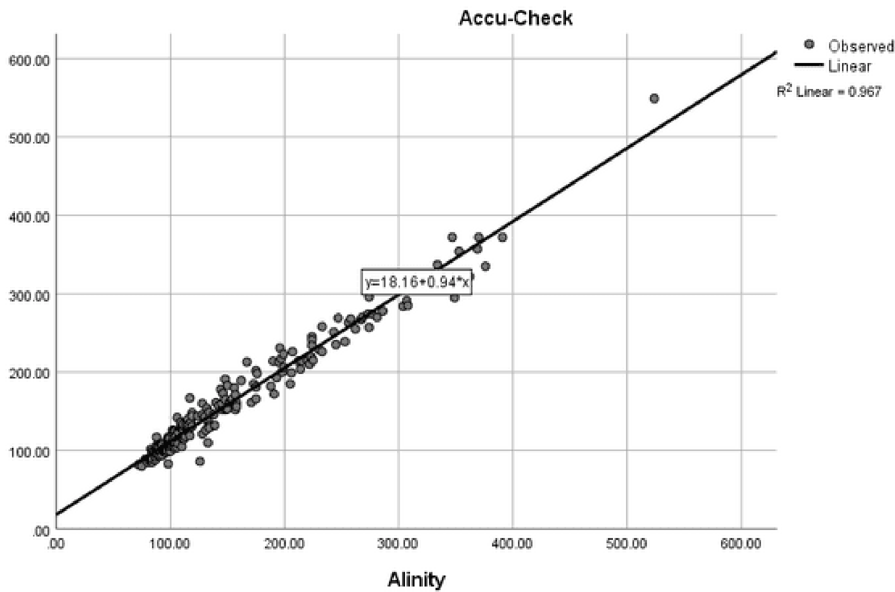
Passing–Bablok regression analysis demonstrated that the Rightest device showed the closest agreement with the reference method, with minimal systematic and proportional bias. In contrast, StatStrip and Accu-Chek Inform II exhibited bothsystematic and proportional deviations, indicating less optimal agreement (Figures 3–5). Detailed regression parameters are provided in the Supplementary Material.

Figure 3: Passing–Bablok regression analysis of StatStrip versus Alinity.



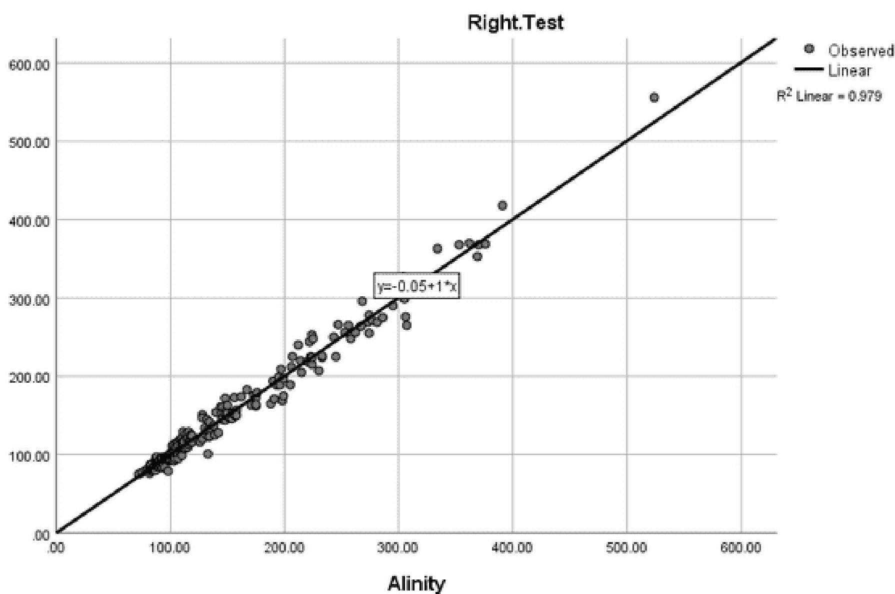
Accu-Chek showed an intercept of 18.158 (95% CI: 13.795 to 22.521) and a slope of 0.935 (95% CI: 0.911 to 0.959),suggesting a greater systematic deviation from the reference method (Figure 4).

Figure 4: Passing–Bablok regression analysis of Accu-Chek versus Alinity.



The Rightest device demonstrated an intercept close to zero (−0.055; 95% CI: −3.735 to 3.625) and a slope of 1.001 (95% CI: 0.981 to 1.022), indicating near-perfect proportional agreement with the Alinity analyzer (Figure 5).

Figure 5: Passing–Bablok regression analysis of Rightest versus Alinity.



Accuracy evaluation based on ISO 15197 criteria demonstrated variability in analytical performance among the devices (Table 3). The Rightest device achieved the highest overall compliance, with accuracy of 98.0% at glucose levels <100 mg/dL, 94.9% at 100–200 mg/dL, and 100% at >200 mg/dL. StatStrip showed good performance at low and high glucose ranges but reducedcompliance in the intermediate range (83.7%). Accu-Chek Inform II demonstrated lower accuracy at glucose levels <200 mg/dL but high compliance at levels >200 mg/dL (98.0%). Additional descriptive statistics, detailed ISO analysis, and full regression results are provided in the Supplementary Material.

Table 3: Accuracy of POCT devices based on ISO 15197 criteria.

Device	<100 mg/dL (%)	100–200 mg/dL (%)	>200 mg/dL (%)
StatStrip	98.0	83.7	98.0
Accu-Chek	65.3	67.3	98.0
Rightest	98.0	94.9	100

Discussion

This study provides a comprehensive evaluation of the analytical performance of three point-of-care testing (POCT) glucose systems using a standardized laboratory reference method. While all evaluated devices demonstrated strong correlations with the hexokinase-based reference, notable differences in analytical accuracy and bias were observed, highlighting the inherent variability among POCT technologies [10,13,21].

The observed variability in performance is consistent with previous studies demonstrating that device-specific analytical characteristics are influenced by the underlying enzymatic detection principles. Systems based on glucose dehydrogenase (GDH) and its variants are known to be susceptible to certain analytical interferences, whereas glucose oxidase-based systems may be affected by oxygen concentration and sample conditions [23,25]. These methodological differences likely contribute to the variation in agreement observed across devices in this study.

Importantly, the findings underscore that correlation alone does not equate to analytical interchangeability. Despite strong linear associations, clinically relevant biases may persist, particularly at extreme glucose concentrations [10,13]. This has direct implications for patient management, as even small deviations may influence clinical decision-making in glycemic control.

The selection of outpatient participants in this study allowed for a more controlled evaluation of device performance by minimizing confounding factors such as acute physiological instability, medication effects, and critical illness–related alterations commonly encountered in emergency and intensive care settings [22,24]. However, this also highlights the need for caution when extrapolating these findings to inpatient or critically ill populations, where additional variables may further impact measurement accuracy.

POCT glucose devices are widely used across both outpatient and inpatient settings due to their rapid turnaround time and operational convenience [9,24]. Nevertheless, their analytical performance may vary depending on patient condition, particularly in critically ill individuals where factors such as impaired perfusion, altered hematocrit, and oxygenation status can significantly influence results [17,19].

Several known interferences may affect the accuracy of POCT glucose measurements. Variations in hematocrit can lead to either underestimation or overestimation of glucose levels depending on the device technology. Additionally, the presence of reducing substances such as ascorbic acid, galactose, and other endogenous or exogenous compounds may interfere with electrochemical detection systems [2,23]. Oxygen tension and sample handling conditions may further contribute to measurement variability [7]. These factors collectively emphasize the importance of understanding device-specific limitations in clinical practice.

Despite the demonstrated analytical reliability of POCT systems, it is important to emphasize that these devices are not intended for the diagnosis of diabetes. Diagnostic criteria require the use of standardized laboratory-based methods with established traceability and analytical performance [3,20]. POCT glucose measurements are primarily intended for monitoring and immediate clinical decision-making rather than definitive diagnosis.

Overall, these findings reinforce the importance of careful device selection, ongoing quality assurance, and appropriate clinical interpretation when using POCT glucose systems. Integration of POCT technologies with laboratory standards remains essential to ensure accurate and reliable patient care across diverse healthcare settings [21,27].

Conclusion

This study provides a comparative evaluation of three point-of-care glucose testing (POCT) systems StatStrip, Accu-Chek Inform II, and Rightest against a hexokinase-based laboratory reference method using the Alinity C platform. All evaluated devices demonstrated strong correlations with the reference method, confirming their utility for glucose monitoring in routine clinical practice.

However, clinically relevant differences in analytical agreement and accuracy were observed. Among the evaluated systems, Rightest demonstrated the closest agreement with the reference method, with minimal systematic and proportional bias, whereas StatStrip and Accu-Chek Inform II showed greater analytical variability. These findings highlight that high correlation does not necessarily equate to analytical agreement, emphasizing the importance of comprehensive performance evaluation beyond correlation metrics alone.

Importantly, variability in accuracy across glucose ranges—particularly in intermediate concentrations - suggests that device performance may influence clinical decision-making in specific scenarios. Therefore, POCT glucose systems should be used with an understanding of their analytical limitations and are best positioned as tools for monitoring rather than definitive diagnostic instruments.

Overall, careful device selection, implementation of rigorous quality assurance programs, and appropriate clinical interpretation are essential to ensure safe and effective use of POCT glucose testing. Strengthening the integration between POCT systems and standardized laboratory methods may further improve diagnostic reliability and support optimal glycemic management across diverse healthcare settings.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Access to the data is subject to ethical approval and participant confidentiality regulations.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest related to this study.

Funding

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

Acknowledgements

The authors would like to thank the laboratory staff and clinical team for their assistance in data collection and technical support. The authors also acknowledge the affiliated institution and hospital for providing facilities and administrative support.

Author Contributions

All authors contributed to the study conception and design. Data collection and analysis were performed by the authors. The first draft of the manuscript was written by the first author, and all authors critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

References

1. Al-Hasani W, Ranasinghe R, Rogers H, Spanier W, Spears K, Gayle C, et al. Clinical utility of point-of-care glucose testing in the assessment of gestational diabetes: a prospective cohort study. BJOG. 2024;131(9):1270-1278. doi:10.1111/1471-0528.17811
2. Albloui F, John J, Alghamdi O, Alseraye F, Alqahtani A, Tamimi W, et al. Effect of hematocrit, galactose and ascorbic acid on blood glucose readings of three point-of-care glucometers. Scand J Clin Lab Invest. 2022;82(7-8):563-570. doi:10.1080/00365513.2022.2138779
3. American Diabetes Association Professional Practice Committee. Introduction: Standards of Medical Care in Diabetes-2022. Diabetes Care. 2022;45(Suppl 1):S1-S2. doi: doi:10.2337/dc22-S001
4. Baumstark A, Jendrike N, Kamecke U, Liebing C, Pleus S, Freckmann G. Measurement accuracy of two professional-use systems for point-of-care testing of blood glucose. Clin Chem Lab Med. 2020;58(3):445-455. doi:10.1515/cclm-2019-0686
5. Ba Y, Xu J, Yuan L, Zhu H, Yang Y, Lam MM, et al. Assessment of the performance of blood glucose monitoring systems for monitoring dysglycaemia in neonatal patients. BMJ Paediatr Open. 2018;2(1):e000339. doi:10.1136/bmjpo-2018-000339
6. Bhansali D, Chima HS, Peretti AR, Ramarajan V. Comparative testing for better glycemic control. Lab Med. 2009;40(8):478-481. doi:10.1309/LMYZA3ICR7OMRCP5
7. Bruns DE, Knowler WC. Stabilization of glucose in blood samples: why it matters. Clin Chem. 2009;55(5):850-852. doi:10.1373/clinchem.2008.116564
8. Chen H, Yao Q, Dong Y, Tang Z, Li R, Cai B, et al. The accuracy evaluation of four blood glucose monitoring systems according to ISO 15197 criteria. Prim Care Diabetes. 2019;13(3):252-258. doi:10.1016/j.pcd.2019.03.009
9. El-Osta A, Woringer M, Pizzo E, Verhoef T, Dickie C, Ni MZ, et al. Does use of point-of-care testing improve cost-effectiveness of the NHS Health Check programme in primary care? BMJ Open. 2017;7(8):e015494. doi:10.1136/bmjopen-2016-015494
10. Ekhlaspour L, Mondesir D, Lautsch N, Balliro C, Hillard M, Magyar K, et al. Comparative accuracy of 17 point-of-care glucose meters. J Diabetes Sci Technol. 2017;11(3):558-566. doi:10.1177/1932296817694773
11. Fabre-Estremera B, Martínez-Chávez E, Manzano Ocaña M, Carcavilla Urquí A, Morales Sánchez MDLA, Pinilla Tejado I, et al. Use of point-of-care glucometers during an oral glucose tolerance test in children for prediabetes and diabetes diagnosis. Adv Lab Med. 2024;5(2):189-196. doi:10.1515/almed-2023-0098
12. Goldstein LN, Wells M, Vincent-Lambert C. The cost-effectiveness of upfront point-of-care testing in the emergency department. Scand J Trauma Resusc Emerg Med. 2019;27(1):110. doi:10.1186/s13049-019-0682-3
13. Khan AI, Vasquez Y, Gray J, Wians FH Jr, Kroll MH. The variability of results between point-of-care testing glucose meters and the central laboratory analyzer. Arch Pathol Lab Med. 2006;130(10):1527-1532. doi:10.5858/2006-130-1527-TVORBP
14. Lockyer MG, Fu K, Edwards RM, Collymore L, Thomas J, Hill T, et al. Evaluation of the Nova StatStrip glucometer in a pediatric hospital setting. Clin Biochem. 2014;47(9):840-843. doi:10.1016/j.clinbiochem.2014.03.003
15. Louie RF, Tang Z, Sutton DV, Lee JH, Kost GJ. Point-of-care glucose testing. Arch Pathol Lab Med. 2000;124(2):257-266. doi:10.1043/0003-9985(2000)124<0257:POGT>2.0.CO;2
16. Mayer-Davis EJ, Kahkoska AR, Jefferies C, Dabelea D, Balde N, Gong CX, et al. ISPAD Clinical Practice Consensus Guidelines 2018: Definition, epidemiology, and classification of diabetes in children and adolescents. Pediatr Diabetes. 2018;19(Suppl 27):7-19. doi:10.1111/pedi.12773
17. Nichols JH, Brandler ES, Fantz CR, Fisher K, Goodman MD, Headden G, et al. A multicenter evaluation of a point-of-care blood glucose meter system in critically ill patients. J Appl Lab Med. 2021;6(4):820-833. doi:10.1093/jalm/jfaa128
18. Sönmez Ç, Babacan AD, Akkaya N, Kaymak AO, Sasmaz G, Akın KO. Evaluation of Accu-Chek Inform II and hexokinase method in emergency departments for glucose measurement. Gazi Med J. 2015;26(2):72-77. (no DOI available)
19. Rebel A, Rice MA, Fahy BG. The accuracy of point-of-care glucose measurements. J Diabetes Sci Technol. 2012;6(2):396-411. doi:10.1177/193229681200600233
20. Sacks DB, Arnold M, Bakris GL, Bruns DE, Horvath AR, Lernmark Å, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. Diabetes Care. 2023;46(10):e151-e199. doi:10.2337/dc23-0398 2023;60(4):290-303. doi:10.1080/104

08363.2023.2170316

21. Pleus S, Jendrike N, Freckmann G. Accuracy evaluation of blood glucose monitoring systems according to ISO 15197:2013. Diabetes Technol Ther. 2020;22(5):362-370. doi:10.1089/dia.2019.0406
22. Nichols JH. Point-of-care testing in critical care medicine. Clin Lab Med. 2021;41(1):1-17. doi:10.1016/j.cll.2020.10.001
23. Tang Z, Louie RF, Lee JH, Kost GJ. Effects of hematocrit and interfering substances on glucose meter performance. Clin Chim Acta. 2021;520:7-14. doi:10.1016/j.cca.2021.05.010
24. Tran NK, Godwin ZR, Baird GS. Clinical impact of point-of-care glucose testing in hospital practice. Crit Rev Clin Lab Sci.
25. Freckmann G, Pleus S, Link M. Accuracy of blood glucose monitoring systems: current status and future perspectives. J Diabetes Sci Technol. 2022;16(5):1106-1114. doi:10.1177/19322968211006760
26. Jendrike N, Baumstark A, Freckmann G. Evaluation of analytical performance of glucose monitoring systems in clinical settings. Diabetes Ther. 2021;12(2):545-557. doi:10.1007/s13300-020-00963-0
27. Klonoff DC, Ahn D, Drincic A. Continuous and point-of-care glucose monitoring in hospital settings. J Diabetes Sci Technol. 2021;15(5):1036-1045. doi:10.1177/1932296821999064

Supplementary Tables

Supplementary Table 1: Descriptive statistics of blood glucose measurements obtained from POCT devices and the reference laboratory method.

Device	Median (Range)	Mean + SD	CV
Alinity	128 (72 - 524)	159.541 + 81.209	0.509
StatStrip	135.5 (75 - 520)	162.434 + 77.479	0.477
Accu-Chek	141 (80 - 549)	167.337 + 77.212	0.461
Rightest	125.5 (75 - 556)	159.679 + 82.158	0.515

Abbreviations: SD, standard deviation; CV, coefficient of variation.

Supplementary Table 2: Descriptive analysis of differences (delta values) between POCT devices and the reference method.

Comparison	Median (Range)	Mean + SD	CV
Alinity vs StatStrip	-3 (-48 to 43)	-2.89 + 13.63	-4.71
Alinity vs Accu-Chek	-8 (-50 to 54)	-7.80 + 14.95	-1.92
Alinity vs Rightest	0.5 (-32 to 42)	-0.14+ 11.80	-85.59

Negative values indicate higher glucose readings from POCT devices compared to the reference method.

Supplementary Table 3: Normality test of glucose measurements using the Kolmogorov–Smirnov test (n = 196).

Device	Kolmogorov-Smirnov Value	p-value
Alinity	0.171	<0.001
StatStrip	0.150	<0.001
Accu-Chek	0.145	<0.001
Rightest	0.171	<0.001

All datasets showed non-normal distribution ($p < 0.05$), supporting the use of non-parametric statistical analysis.

***indicates strong correlation. Spearman's correlation was used due to non-normal data distribution.*

Supplementary Table 4: Correlation analysis between POCT devices and the reference method.

Comparison	Spearman's r	p-value
StatStrip vs Alinity	0.975**	<0.001
Accu-Chek vs Alinity	0.967**	<0.001
Rightest vs Alinity	0.982**	<0.001

Supplementary Table 5: Comparative statistical analysis between POCT devices and the reference method.

Comparison	Median (Range)	Mean	CV	Median	Mean	CV	p-value
StatStrip vs Alinity	128 (72-524)	159.541 ± 81.21	0.509	135.5 (75-520)	162.434 ± 77.48	0.477	<0.001
Accu-Chek vs Alinity	128 (72-524)	159.541 ± 81.21	0.509	141 (80-549)	167.337 + 77.21	0.461	<0.001
Rightest vs Alinity	128 (72-524)	159.541 + 81.21	0.509	125.5 (75–556)	159.679 + 82.16	0.515	<0.001

p-values derived from non-parametric comparison tests.

Supplementary Table 6: Analytical accuracy of POCT devices according to ISO 15197:2013 criteria across glucose concentration ranges.

Category	<100 mg/dL	100–200 mg/dL	>200 mg/dL	Total
Not compliant	1 (2.0%)	16 (16.3%)	1 (2.0%)	18 (9.2%)
Compliant	48 (98.0%)	82 (83.7%)	48 (98.0%)	178 (90.8%)

Category	<100 mg/dL	100–200 mg/dL	>200 mg/dL	Total
Not compliant	17 (34.7%)	32 (32.7%)	1 (2.0%)	50 (25.5%)
Compliant	32 (65.3%)	66 (67.3%)	48 (98.0%)	146 (74.5%)

Category	<100 mg/dL	100–200 mg/dL	>200 mg/dL	Total
Not compliant	1 (2.0%)	5 (5.1%)	0 (0.0%)	6 (3.1%)
Compliant	48 (98.0%)	93 (94.9%)	49 (100%)	190 (96.9%)

StatStrip / Accu-Chek / Rightest

Accuracy was evaluated according to ISO 15197:2013 criteria, which require that ≥95% of results fall within ±15 mg/dL (forglucose <100 mg/dL) or ±15% (for glucose ≥100 mg/dL) of the reference method.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research article

Autoantibody Profiles and Clinical Correlations in Systemic Sclerosis: A Cross-Sectional Study

Debamita Datta[2], Sushil Kumar[1*], Lekha Priyadharshini Kamarajan[1], Mala Mahto[1], Bandana Kumari[1], Ayan Banerjee[1], Swetalina Pradhan[3], Bijay Nanda Naik[4]

[1]*Department of Biochemistry, All India Institute of Medical Sciences (AIIMS), Patna, Bihar, India*
[2]*Department of Biochemistry, Tripura Santiniketan Medical College, Agartala, Tripura, India*
[3]*Department of Dermatology, All India Institute of Medical Sciences (AIIMS), Patna, Bihar, India*
[4]*Department of Community and Family Medicine, All India Institute of Medical Sciences (AIIMS), Patna, Bihar, India*

Article Info

****Corresponding Author:***
Sushil Kumar
All India Institute of Medical Sciences (AIIMS), Patna, Bihar, India
E-mail: drsushilk@aiimspatna.org

Keywords

Systemic sclerosis, Antinuclear antibodies, Anti-Scl-70, Interstitial lung disease

Abstract

Background: Systemic sclerosis (SSc) is a rare autoimmune disease marked by immune dysregulation, vascular injury, and fibrosis. Antinuclear antibodies (ANA) are pivotal for diagnosis and prognosis, yet Indian data remain scarce. This study assessed ANA patterns and autoantibody profiles in SSc and their clinical relevance.

Methods: A cross-sectional study was conducted in the Department of Dermatology, AIIMS Patna, from September 2022 to September 2023. Eighty-five patients fulfilling ACR/EULAR criteria for SSc were enrolled. Clinical features were documented using a structured proforma, and patients were classified into diffuse cutaneous (dcSSc) and limited cutaneous (lcSSc) subsets. ANA screening was performed by indirect immunofluorescence (IIF) on HEp2 cells, followed by autoantibody profiling with line immunoassay (LIA). Statistical analyses were performed using JAMOVI v2.3.28, with $p<0.05$ considered significant.

Results: Of 85 patients, 70 (82.4%) were female (ratio of 5:1), with a mean age of 34.7 ± 10.7 years. dcSSc accounted for 55.3%. Common manifestations included arthralgia (69.4%), respiratory symptoms (68.2%), and interstitial lung disease (ILD) (64.1%). Homogeneous ANA was the predominant IIF pattern (41.2%), followed by centromere (10.6%) and speckled (8.2%). LIA revealed anti-Scl70 as the most frequent autoantibody (58.3%), significantly associated with ILD and Raynaud's phenomenon. Anti-U1 snRNP and anti-PMScl antibodies correlated with musculoskeletal involvement.

Conclusion: SSc in this cohort showed female predominance with frequent arthralgia and respiratory symptoms. Anti-Scl-70 antibodies were most common, strongly linked to diffuse disease and ILD. ANA profiling provides disease characterization in Indian SSc.

Introduction

Systemic sclerosis (SSc) is a heterogeneous, uncommon autoimmune connective tissue disease characterized by a triad of immune dysregulation, widespread microvascular injury, and progressive fibrosis affecting the skin and internal organs [1,2]. The estimated prevalence in India is approximately 12 per 100,000 population, with a striking female predominance [1-3] and a reported female-male ratio ranging from 3:1 to 8:1 [3,4]. The pathogenesis of systemic sclerosis is complex and involves endothelial dysfunction, chronic inflammation, and aberrant fibroblast activation, leading to excessive extracellular matrix deposition. Clinically, SSc affects not only the skin but also several internal organs, including the lungs, heart, kidneys, gastrointestinal tract, and musculoskeletal system, resulting in diverse clinical manifestations and variable disease outcomes [2,5,6]. Based on the extent of cutaneous involvement, SSc is classified into two major subsets: (lcSSc) and dcSSc [4,7]. lcSSc is characterized by fibrosis confined to the acral regions, including the face and distal extremities, whereas dcSSc involves the trunk and proximal limbs and is frequently associated with earlier and more severe internal organ involvement [1,7]. Autoantibodies are a hallmark of SSc and play a pivotal role in diagnosis, prognosis, and disease stratification. ANA are the most frequently detected [2,8]. IIF on HEp2 cells is considered the gold standard for screening [8] because of its high sensitivity and ability to identify distinct staining patterns. Enzyme-linked immunosorbent assay (ELISA), although more specific, allows detection and quantification of individual autoantibodies; however, in recent years, LIA has emerged as a reliable and efficient technique for ANA profiling. This method enables simultaneous detection of multiple nuclear and cytoplasmic antigens on a single nitrocellulose strip, offering faster processing time with diagnostic performance comparable to ELISA [9]. Given the established associations between specific autoantibodies and distinct clinical phenotypes, a comprehensive autoantibody profiling provides valuable insights into disease stratification, prognosis, and clinical management of patients with systemic sclerosis [2,10]. Among the autoantibodies, anti-topoisomerase I (Anti Scl70), anti-centromere, and anti-PMScl are particularly relevant, as they correlate with distinct clinical phenotypes and organ involvement. For instance, Anti Scl70 is strongly associated with ILD [5,6], while anti-centromere antibodies are more frequent in lcSSc and linked to vascular manifestations [4]. Despite extensive research in Western populations, data from Indian cohorts remain limited. Regional variations in clinical presentation and autoantibody distribution highlight the importance of population-specific studies [10]. This study was therefore undertaken to evaluate the clinical features, ANA patterns, and autoantibody profiles of SSc patients from Bihar and to explore their associations with disease subtypes and organ involvement.

Materials and Methods

Study Design and Setting

This was an observational cross-sectional study conducted in the Department of Dermatology and the Department of Biochemistryat All India Institute of Medical Sciences (AIIMS), Patna, Bihar, India, between September 2022 and September 2023. All patients presenting to the outpatient and inpatient services with a diagnosis of SSc during the study period were approached for inclusion.

Study Population

A total of 85 adult patients meeting the American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria for SSc were enrolled [7,11]. The following groups were excluded during patient selection:

Individuals with overlap connective tissue diseases Individuals with other autoimmune rheumatic disorders lacking features of systemic sclerosis

Clinical Evaluation

Demographic details, disease duration, and clinical features were recorded using a structured proforma. Organ involvement was assessed through clinical examination and relevant investigations, including high-resolution computed tomography (HRCT) of the thorax for ILD [6], urine analysis for proteinuria, and gastrointestinal symptom evaluation.

Clinical Classification

Patients were further categorized into dcSSc and lcSSc subtypes based on the extent of skin involvement. dcSSc was defined by the rapid progression of skin thickening from fingers to trunk, often accompanied by severe constitutional symptoms, pulmonary fibrosis, and cardiac or renal involvement. lcSSc was characterized by long-standing Raynaud's phenomenon, limited skin thickening confined to the face and distal extremities, and milder systemic features. Patients without skin thickening were also included in the lcSSc group.

Autoantibody Testing

Serum samples of all the selected patients underwent ANA screening by IIF, subsequent ANA profiling by using LIA.
For IIF, serum samples were tested using the EUROIMMUN biochip technique (Germany), which employs HEp2 cells and monkey liver tissue sections. Samples were diluted 1:100 in phosphate-buffered saline and incubated with fluorescein-labelled anti-human globulin. Fluorescence was visualized under an inverted microscope at 10× and 40× magnification, with intensity compared against positive and negative controls [12].
For autoantibody profiling, LIA was performed using commercially available (human diagnostics, made in India, in collaboration with Germany) strips coated with nuclear and cytoplasmic antigens. Patient sera were incubated with the strips, and bound antibodies were detected using horseradish peroxidase (HRP)labeled antihuman IgG. The presence of specific autoantibodies (including anti-Scl70, anti-centromere (CENPB), anti-PMScl, anti-Ku, and anti-U1snRNP) was visualized as brown bands following the addition of substrate and stop solutions, as per accordance with the manufacturer's protocol [13].

Statistical Analysis
Data were entered into Microsoft Excel and analyzed using JAMOVI version 2.3.28. Continuous variables were expressed as mean ± standard deviation (SD). Between group comparisons for continuous variables were performed using independent samples 't' test. Categorical variables were compared using Chisquare test. A p-value <0.05 was considered statistically significant.

Results
Demographic and Clinical Characteristics
A total of 85 systemic sclerosis (SSc) patients were enrolled. The mean age at evaluation was 34.7 ± 10.7 years, with a meandisease duration of 43.7 ± 35.4 months. There was a marked female predominance (82.4%; femaletomale ratio 5:1). Diffuse cutaneous systemic sclerosis was observed in 47 patients (55.3%), while 38 patients (44.7%) had limited cutaneous systemic sclerosis.

Clinical Features
The most frequent clinical manifestations included joint pain (69.4%), respiratory symptoms (68.2%), and ILD confirmed on HRCT thorax (64.1%). Skin thickening was observed in 55.3% of patients, gastrointestinal symptoms in 42.4%, Raynaud's phenomenon in 38.8%, and proteinuria in 27.1% (Table 1).

Table 1: Clinical features of Systemic Sclerosis patients (n=85).

Clinical Feature	Frequency (n)	Percentage (%)
Joint pain	59	69.4
Respiratory symptoms	58	68.2
Interstitial lung disease	55	64.1
Skin thickening	47	55.3
Gastrointestinal symptoms	36	42.4
Raynaud's phenomenon	33	38.8
Proteinuria	23	27.1

Values are presented as numbers (percentage). HRCT= high-resolution computed tomography.

ANA Patterns by IIF
The distribution of ANA patterns detected by IIF according to disease subtype and gender is presented in Table 2. ANA screening revealed that the homogeneous pattern was the most common (41.2%), followed by centromere (10.6%), speckled(8.2%), and nucleolar (7.0%) patterns. Mixed ANA patterns were observed in 15.2% of patients, while 17.6% tested negative by IIF. The homogenous pattern was distributed across both dcSSc and lcSSc subtypes and was more frequent in females.

Table 2: ANA patterns by IIF and distribution by disease subtype and gender.

ANA patterns by IIF	Total (%)	Types of SSc	Types of SSc	Gender	Gender
ANA patterns by IIF	Total (%)	Diffuse	Limited	Female	Male
Homogeneous	35 (41.1%)	24	11	29	6
Nucleolar	6 (7.0%)	4	2	6	0
Speckled	6 (8.2%)	4	2	5	1
Centromere	9 (10.6%)	3	6	7	2
Mixed Homogeneous -Nucleolar	5 (4.7%)	2	3	4	1
Mixed Homogeneous -Speckled	5 (5.9%)	1	4	3	2
Mixed Homogeneous -Nucleolar-Nuclear Membrane	2 (2.3%)	1	1	2	0
Mixed Speckled-Cyt oplasmic	2 (2.3%)	1	1	1	1

ANA patterns by IIF	Total (%)	Types of SSc	Types of SSc	Gender	Gender
Negative	15 (17.6%)	7	8	13	2

Values are presented as a number (percentage). ANA= antinuclear antibody; IIF= indirect immunofluorescence; dcSSc= diffuse cutaneous systemic sclerosis; lcSSc= limited cutaneous systemic sclerosis

Organ Involvement and ANA Associations

The association between ANA patterns detected by IIF and organ system involvement is summarized in Table 3. Homogenous ANA was strongly associated with pulmonary disease (47.4%), while centromere antibodies correlated with renal and hepaticinvolvement. Negative ANA cases also demonstrated organ involvement, particularly pulmonary manifestations (Table 3). These associations reinforce the prognostic role of ANA patterns in systemic sclerosis.

Table 3: Association between ANA patterns and organ involvement.

ANA Pattern	Pulmonary n (%)	Musculoskeletal n (%)	Renal n (%)	Hepatic n (%)	CNS n (%)
Homogeneous	27(47.4)	4(30.8)	4(50.0)	0(0)	1(100)
Nucleolar	6(10.5)	1(7.7)	0(0)	0(0)	0(0)
Speckled	5(8.8)	1(7.7)	1(16.7)	0(0)	0(0)
Centromere	5(8.8)	0(0)	1(16.7)	1(50.0)	0(0)
Mixed homogeneous-nucleolar	3(5.3)	1(7.7)	0(0)	0(0)	0(0)
Mixed homogenous-speckled	2(3.5)	1(7.7)	0(0)	0(0)	0(0)
Mixed Homogeneous -Nucleolar-Nuclear membrane	1(1.7)	1(7.7)	1(16.7)	1(50.0)	0(0)
Mixed speckled and cytoplasmic	2(3.5)	0(0)	0(0)	0(0)	0(0)
ANA Negative	6(10.5)	4(30.8)	0(0)	0(0)	0(0)
Total	57(100)	13(100)	6(100)	2(100)	1(100)

Values are presented as percentages. IIF=indirect immunofluorescence; CNS= central nervous system.

Autoantibody Profiles by LIA

The distribution of autoantibody profiles according to disease subtype and gender is shown in Table 4. Autoantibody profiling using LIA yielded positive results in 84 patients, while one patient tested negative despite having clinical features consistent with systemic sclerosis.

Anti-Scl70 was the most frequent autoantibody, detected in 58.3% of patients. It was more common in dcSSc (61.7%) compared to lcSSc (54.1%) and was predominantly observed in females. Other autoantibodies included anti-PMScl (8.3%), anti-CENPB (6.0%), anti-Ku (3.6%), and anti-U1snRNP (3.6%). Mixed antibody profiles were identified in 20.2% of patients (Table 4).

Table 4: Distribution of autoantibody profiles according to disease subtype and gender.

Autoantibody	Total (%)	dcSSc (%)	lcSSc (%)	Female (%)	Male (%)
Anti-Scl-70	49 (58.3)	29 (61.7)	20 (54.1)	40 (58.0)	9 (60.0)
Anti-PM-scl	7 (8.3)	6 (12.8)	1 (2.7)	6 (8.7)	1 (6.7)
Anti-CENP-B	5 (6.0)	2 (4.3)	3 (8.1)	4 (5.8)	1 (6.7)
Anti-Ku	3 (3.6)	2 (4.3)	1 (2.7)	3 (4.3)	0
Anti-U1-snRNP	3 (3.6)	2 (4.3)	1 (2.7)	2 (2.9)	1 (6.7)
Mixed	17 (20.2)	6 (12.8)	11 (29.8)	14 (20.2)	3 (20)

Autoantibody	Total (%)	dcSSc (%)	lcSSc (%)	Female (%)	Male (%)
Total	(n=84) 100%	(n=47) 100%	(n=37) 100%	(n=69) 100%	(n=15) 100%

Values are presented as a number (percentage). LIA=line immunoassay; dcSSc=diffuse cutaneous systemic sclerosis; lcSSc=limited cutaneous systemic sclerosis.

Autoantibodies and Organ System Involvement

The association between autoantibodies detected by LIA and organ system involvement is summarized in Table 5. Anti-Scl-70 antibodies were most frequently associated with pulmonaryinvolvement, followed by musculoskeletal and renal involvement. Other autoantibodies, including anti-U1-snRNP, anti-PM-Scl, and anti-CENP-B, were observed less frequently and across multiple organ systems.

Table 5: Association of Autoantibodies with Organ Involvement.

Organ System	Anti-Scl-70 n (%)	Anti-U1-snRNP n (%)	Anti-PM-Scl n (%)	Anti-CENP-B n (%)	Other/Mixed n (%)
Pulmonary (n = 57)	35 (61.4)	5 (8.8)	2 (3.5)	2 (3.5)	13 (22.8)
MKS (n = 13)	6 (46.2)	1 (7.7)	1 (7.7)	0	5 (38.4)
GIT (n = 2)	1 (50.0)	0	0	0	1 (50.0)
Renal (n = 6)	3 (50.0)	1 (16.7)	0	1 (16.7)	1 (16.7)
CVS (n = 2)	1 (50.0)	0	0	0	1 (50.0)
Hepatic (n = 2)	0	0	1 (50.0)	0	1 (50.0)
CNS (n = 1)	0	1 (100)	0	0	0

Values are presented as a number (percentage). MKS=Musculoskeletal, GIT= Gastrointestinal system, CVS= cardiovascular system, CNS= Central nervous system

Association Between Autoantibodies and Specific Clinical Features

The association between specific autoantibodies and selected clinical features is presented in Table 6. A statistically significant association was observed between anti-Scl-70 antibodies andRaynaud's phenomenon (p = 0.002) as well as interstitial lung disease (p = 0.001). Anti-U1snRNP and anti-PM-scl were inversely linked to joint pain (p=0.024) and Raynaud's phenomenon (p=0.005).

Table 6: Clinical associations of specific autoantibodies in Systemic Sclerosis *(n=85)*.

Antibody	Character	Negative (N/n)	Positive (F/f)	p
Anti-Scl-70	Raynaud's phenomenon	3/24	30/61	0.002
	ILD	9/24	46/61	0.001
	Joint pain	15/24	44/61	0.386
Anti-U1-snRNP	Raynaud's phenomenon	30/73	3/12	0.005
	ILD	48/73	7/12	0.618
	Joint pain	54/73	5/12	0.024
Anti-PM-scl	Raynaud's phenomenon	33/74	0/11	0.005
	ILD	50/74	5/11	0.152
	Joint Pain	51/74	8/11	0.798

Values are numbers. The table demonstrates the distribution of selected clinical features among patients positive and negative for Scl70, U1snRNP, and PMScl auto-antibodies. ILD= Interstitial lung disease. Values are presented as numbers. p-values < 0.05 were considered statistically significant.
N = antibodynegative patients with the clinical feature; n = total antibodynegative patients; F = antibodypositive patients with the clinical feature; f = total antibodypositive patients

Discussion

Systemic sclerosis is a heterogeneous autoimmune disorder with variable clinical manifestations and immunological profiles. In the present study, a marked female predominance and mean age of onset (34.7 years) were observed, with females accounting for 82.4% of the study population and a female-to-male ratio of 5:1. This finding is consistent with previous Indian studies and it reflects the well-recognised female preponderance of SSc reported globally [1]. The mean age of patients in our cohort was comparable to that reported by Sharma et al., further supporting demographic similarities within Indian populations [1]. dcSSc was more frequently observed than the limited cutaneous subtype in the present study. This distribution is broadly comparable to the findings of Pradhan et al. in a western Indian cohort, although slight differences in proportions may reflect regional variations, referral patterns, and differences in sample size [2]. The higher prevalence of dcSSc in our cohort may also explain the increased frequency of internal organ involvement observed.

The most common clinical manifestations in our study were joint pain, respiratory symptoms, and skin thickening. The high prevalence of joint pain is consistent with previous reports [14-16], although variability in reported musculoskeletal involvement across studies has been attributed to differences in diagnostic criteria and symptom definitions [5,10,11,17]. Respiratory involvement was common in our cohort, with ILD detected in a substantial proportion of patients. Although the frequency of pulmonary involvement in our study was higher than that reported in some large longitudinal cohorts, this discrepancy may be explained by differences in study duration, sample size, and the use of high-resolution computed tomography for detection [5,6,10].

ANA testing by IIF demonstrated that the homogeneous pattern was the most frequently observed ANA pattern in our cohort. This finding is in agreement with prior studies that have highlighted the predominance of homogeneous and nucleolar patterns in systemic sclerosis [2]. The presence of ANA negativity in a subset of patients highlights the clinical heterogeneity of SSc and highlights the importance of comprehensive immunological evaluation. Autoantibody profiling by line immunoassay revealed anti-Scl-70 as the most prevalent autoantibody, particularly among patients with diffuse cutaneous disease and females. This finding aligns with previous studies demonstrating a strong association between anti-Scl-70 antibodies and dcSSc, as well as severe internal organ involvement, especially pulmonary involvement and Raynaud's phenomenon [5]. The observed association between anti-Scl-70 antibodies and interstitial lung disease in our study further reinforces the prognostic significance of this autoantibody [4,5]. Given the established role of autoantibodies in the diagnosis and prognostic stratification of SSc, comprehensive immunological profiling was undertaken. Indirect immunofluorescence (IIF) on HEp-2 cells remains the reference method and gold standard for antinuclear antibody (ANA) screening in systemic autoimmunerheumatic diseases, including SSc [8,17]. Large cohort studies have reported ANA positivity in approximately 85–95% of SSc patients [18,19], rendering true ANA-negative disease relatively uncommon. In the present study, however, 17.6% of patients were ANA-negative by IIF. This comparatively higher rate may be attributed to methodological and immunological factors. Certain SSc-associated autoantibodies may produce weak, low-titre, or atypical fluorescence patterns that are difficult to interpret, particularly at higher screening dilutions, thereby increasing the likelihood of under-detection [20]. Additionally, variability in antigen expression on HEp-2 substrates and technical aspects of assay performance may further influence sensitivity. Patients with "Seronegative" SSc (negative on IIF) may possess rare or novel antibodies, such as anti-U11/U12 RNP, which can be detected by specialized immunoblot assays [21]. Pulmonary involvement emerged as the most frequently affected organ system in the present cohort, with anti-Scl-70 antibodies showing the strongest association. Similar variations in autoantibody distribution and clinical correlations have been reported in international cohorts from New Zealand, Brazil, and Poland, where distinct antibody profiles were observed across different populations [22–24]. These findings highlight the influence of geographic and ethnic factors on autoantibody expression while reinforcing the prognostic relevance of comprehensive serological profiling. This observation is consistent with earlier reports identifying ILD as a leading cause of morbidity and mortality in SSc [6,14]. Musculoskeletal involvement remained clinically significant due to the high prevalence of joint pain, a finding also supported by previous literature [11,15,16]. Significant inverse associations were also observed between anti-U1-snRNP and Raynaud's phenomenon and joint pain; this indicates that anti- U1snRNP may define a distinct subgroup with musculoskeletal involvement but less vascular disease [25]. Similarly, PMScl antibodies, classically linked to polymyositis–systemic sclerosis overlap, were inversely associated with Raynaud's phenomenon in our cohort, in line with EUSTAR. which demonstrated [25,26] that anti-PMScl antibodies are associated with overlap syndromes, particularly myositis and arthritis. Raynaud's phenomenon and digital ulcers were less frequent in anti-PMScl patients compared to other antibody subsets. ILD was present but showed more favourable functional outcomes than in anti-Scl70-positive patients. Overall, these findings emphasize the value of ANA profiling in systemic sclerosis, not only for diagnosis but also for risk stratification and prognostic assessment. Early identification of high-risk patients, particularly those with Anti-Scl70 antibodies, may facilitate timely monitoring and intervention, thereby improving clinical outcomes. These results contribute to the growing body of evidence supporting the integration of immunological markers into the routine evaluation and management of SSc [10].

Conclusion

This study demonstrates a clear female predominance in SSc and highlights the heterogeneous clinical and immunological profile of the disease. Joint pain and pulmonary involvement were the most common clinical manifestations, with ILD representing thepredominant internal organ involvement. Anti-Scl-70 emerged as the most frequent autoantibody and showed significant associations with dcSSc, ILD, and Raynaud's phenomenon, while anti-U1-snRNP and anti-PM-Scl antibodies were associated with

specific clinical features, with musculoskeletal involvement but less vascular disease. These findings underscore the clinical relevance of comprehensive ANA profiling in SSc and support its role in disease stratification, prognostication, and informed clinical management.

Authors contribution

S.K. and D.D. contributed to the conception and design of the study and were involved in acquisition, analysis, and interpretation of data. **L.P.K., M.M., B.K., and A.B.** drafted the manuscript and revised it critically for important intellectual content. **S.P.** provided clinically relevant samples for the assay. **B.N.N.** contributed to data analysis and interpretation.

All authors have approved the final draft for publication purposes and are accountable for the accuracy and integrity of the work. The corresponding author Dr. Sushil Kumar is responsible for ensuring that the descriptions are accurate and agreed by all author.

Data Availability Statement

The data supporting the findings of this study are not publicly available due to ethical and privacy restrictions but are available from the corresponding author upon reasonable request.

Acknowledgments

The authors would like to acknowledge the Department of Dermatology, AIIMS Patna, for providing clinical support and laboratory facilities required for the conduct of this study. The authors also thank all the patients who consented to participate in this research.

Artificial Intelligence (AI) Use Statement

Artificial intelligence tools (Chat GPT, Copilot) were used for language editing and improving the clarity of the manuscript. The authors take full responsibility for the content, interpretation, and conclusions of the study.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Funding

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

Ethical Approval

The study was approved by the Institutional Ethics Committee of AIIMS Patna (AIIMS/Pat/IEC/PGTh/Jan22/50). Written informed consent was obtained from all participants before enrolment. This study was in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

References

1. Sharma VK, Trilokraj T, Khaitan BK, Krishna SM. Profile of systemic sclerosis in a tertiary care centre in North India. Indian J Dermatol Venereol Leprol. 2006; 72:416–420. doi:PMID: 17179615. accessed: 09/01/2026
2. Pradhan V, Rajadhyaksha A, Nadkar M, Pandit P, Surve P, Lecerf M et al. Clinical and autoimmune profile of scleroderma patients from Western India. Int J Rheumatol. 2014; 2014:983781. doi: 10.1155/2014/983781. Epub 2014 Oct 19. PMID: 25386193; PMCID: PMC4216682. accessed: 09/01/2026
3. Ghosh SK, Bandyopadhyay D, Saha I, Barua JK. Mucocutaneous and demographic features of systemic sclerosis: a profile of 46 patients from eastern India. Indian J Dermatol. 2012;57(3):201–205. doi:10.4103/0019-5154.96193. accessed: 09/01/2026
4. Allanore Y. Limited cutaneous systemic sclerosis: the unfairly neglected subset. J Scleroderma Relat Disord. 2016; 1:241–246 https://doi.org/10.5301/jsrd.5000216 accessed: 09/01/2026
5. Steen VD, Medsger TA Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. Arthritis Rheum. 2000;43(11):2437–2444. doi:10.1002/1529-0131(200011)43:11<2437: AID-ANR10>3.0.CO;2-U accessed: 09/01/2026
6. Cottin V, Brown KK. Interstitial lung disease associated with systemic sclerosis (SSc-ILD). Respir Res. 2019;20(1):13. doi:10.1186/s12931-019-0980-7 accessed: 09/01/2026
7. Johnson SR, van den Hoogen F, Devakandan K, Matucci-Cerinic M, Pope JE. Systemic sclerosis: To subset or not to subset, that is the question. Eur J Rheumatol. 2020;7(Suppl 3): S222–S227. doi:10.5152/eurjrheum.2020.19116. accessed: 09/01/2026
8. Irure-Ventura J, & López-Hoyos M. (2022). The Past, Present, and Future in Antinuclear Antibodies (ANA). Diagnostics (Basel, Switzerland), 12(3), 647. https://doi.org/10.3390/diagnostics12030647 accessed: 09/01/2026
9. Kaur J, Singh J. Comparative Analysis of Line Immunoassay with Immunofluorescence Assay for the Identification of Autoantibodies in Patients with Suspected Systemic Autoimmune Disorders: A Cross-sectional Study. J Clin Diagn Res. 2022;16(4): BC01–BC04. doi:10.7860/JCDR/2022/53126.16448. accessed: 09/01/2026
10. Nihtyanova SI, Sari A, Harvey JC, et al. Using autoantibodies and cutaneous subset to develop outcome-based disease classification in systemic sclerosis. Arthritis Rheum. 2020;72(3):465–476. https://doi.org/10.1002/art.41153 accessed: 09/01/2026
11. Blank RB, Nwawka OK, Yusov AA, Gordon JK. Inflammatory arthritis in systemic sclerosis: What to do? J Scleroderma Relat Disord. 2019;4(1):3–16. doi:10.1177/2397198318779532 accessed: 09/01/2026
12. EUROIMMUN Medizinische Labordiagnostika AG. Anti-nuclear antibodies (ANA) HEp-2 indirect immunofluorescence test (IIFT): instructions for use. Lübeck (DE): EUROIMMUN; 2022. Accessed February 13, 2026.

13. Human Gesellschaft für Biochemica und Diagnostician mbH. ANA profile line immunoassay (LIA): instructions for use. Wiesbaden (DE): Human GmbH; 2022. Accessed February13, 2026
14. Sandler RD, Matucci-Cerinic M, Hughes M. Musculoskeletal hand involvement in systemic sclerosis. Semin Arthritis Rheum. 2020;50(2):329–334. doi: 10.1016/j.semarthrit.2019.11.003 accessed: 09/01/2026
15. Chitale S, Ciapetti A, Hodgson R, et al. Magnetic resonance imaging and musculoskeletal ultrasonography detect andcharacterize covert inflammatory arthropathy in systemic sclerosis patients with arthralgia. Rheumatology. 2010;49(12):2357–2361. https://doi.org/10.1093/rheumatology/keq268 accessed: 09/01/2026
16. Avouac J, Clements PJ, Khanna D, et al. Articular involvement in systemic sclerosis. Rheumatology. 2012;51(8):1347–1356. https://doi.org/10.1093/rheumatology/kes041 accessed: 09/01/2026
17. Agmon-Levin N, Damoiseaux J, Kallenberg C, Sack U, Witte T, Herold M, et al. international recommendations for the assessment of autoantibodies to cellular antigens referred to as anti-nuclear antibodies. Ann Rheum Dis. 2014;73(1):17–23. doi:10.1136/annrheumdis-2013-203863. Accessed February 13, 2026.
18. Ho KT, Reveille JD. The clinical relevance of autoantibodies in scleroderma. Arthritis Res Ther. 2003;5(2):80–93. doi:10.1186/ar628. Accessed February 13, 2026.
19. Meier FM, Frommer KW, Dinser R, et al. Update on the profile of the EUSTAR cohort: an analysis of the EULAR Scleroderma Trials and Research group database. Ann Rheum Dis. 2012;71(8):1355–1360. doi:10.1136/annrheumdis-2011-200742. Accessed February 13, 2026.
20. Mahler M, Fritzler MJ. Epitope specificity and significance in systemic autoimmune diseases. Ann N Y Acad Sci. 2010; 1183:267-287. doi: 10.1111/j.1749-6632.2009.05127. x. PMID: 20146721.Accessed February 13, 2026.
21. Fritzler MJ, Bentow C, Beretta L, Palterer B, Perurena-Prieto J, Sanz-Martínez MT, Guillen-Del-Castillo A, Marín A,Fonollosa-Pla V, Callejas-Moraga E, et al. Anti-U11/U12 Antibodies as a Rare but Important Biomarker in Patients with Systemic Sclerosis: A Narrative Review. Diagnostics. 2023; 13(7):1257. https://doi.org/10.3390/diagnostics13071257
22. Chang WS, Schollum J, White DH, Solanki KK. A cross-sectional study of autoantibody profiles in the Waikato systemic sclerosis cohort, New Zealand. Clin Rheumatol. 2015;34(11):1921-1927. doi: 10.1007/s10067-015-2981-3. Epub 2015 May 29. PMID: 26022620. Accessed February 13, 2026.
23. Müller Cde S, Paiva Edos S, Azevedo VF, Radominski SC, Lima Filho JH. Autoantibody profile and clinical correlation in a group of patients with systemic sclerosis in southern Brazil. Rev Bras Reumatol. 2011;51(4):314-8, 323-324. English, Portuguese. PMID: 21779708.Accessed February 13, 2026.
24. Żebryk P, Przymuszała P, Nowak JK, Piorunek T, Mularek-Kubzdela T, Puszczewicz M. Autoantibodies and clinical correlations in Polish systemic sclerosis patients: a cross-sectional study. J Clin Med. 2023;12(2):657. doi:10.3390/jcm12020657. Accessed February 13, 2026.
25. Hamaguchi Y, Takehara K. Anti-nuclear autoantibodies in systemic sclerosis: News and perspectives. J Scleroderma Relat Disord. 2018;3(3):201–213. doi:10.1177/2397198318783930 accessed: 09/01/2026
26. Lazzaroni MG, Cavazzana I, Colombo E, Dobrota R, Hernandez J, Walker UA, et al. The clinical phenotype and outcome of systemic sclerosis patients with anti-PM/Scl antibodies: Results from the EUSTAR cohort. Rheumatology (Oxford). 2021;60(11):5028-5041. doi:10.1093/rheumatology/keab152. Accessed: 9/01/ 2026.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Advisory and Interpretative Services under ISO 15189:2022: A Prospective Study of Patient-Initiated Laboratory Consultations

Vivek Pant[1*], Santosh Pradhan[1], Keyoor Gautam[2], Devish Pyakurel[2], Abha Shrestha[2], Neha Neupane[2]

[1]*Department of Biochemistry, Samyak Diagnostic Pvt Ltd, Kathmandu, Nepal*
[2]*Department of Pathology, Samyak Diagnostic Pvt Ltd, Kathmandu, Nepal*

Article Info

[*]***Corresponding Author:***
Dr. Vivek Pant
Department of Biochemistry, Samyak Diagnostic Pvt Ltd, Kathmandu, Nepal
E-mail: drvpant@gmail.com

Keywords

Advisory Service; Direct-to-consumer testing; Patient communication; Laboratory consultation; Interpretative comments; Patient satisfaction; Laboratory Physicians; Test Utilization; ISO 15189:2022

Abstract

Background: ISO 15189:2022 requires accredited medical laboratories to establish advisory and interpretative services to support appropriate test selection, result interpretation, and effective utilization of laboratory examinations. Direct-to-consumer and patient-initiated laboratory testing is reshaping traditional healthcare communication. Patients increasingly seek direct consultation with laboratory physicians for test selection and result interpretation. This study describes the drivers, patterns and perceived value of such interactions in routine laboratory practice.

Methods: A prospective observational study was conducted over 12 months at a high-volume ISO 15189:2022 accredited referral laboratory. All unsolicited, patient-initiated advisory consultations with laboratory physicians were systematically documented using a structured logbook. Consultations were categorized by type (pre-examination advice, post-examination interpretation, clarification of abnormal findings, and resolution of inter-laboratory discrepancies). All abnormal or discordant results discussed during consultation had undergone analytical verification in accordance with internal quality control and validation procedures. Patient satisfaction was assessed using a 5-point Likert scale. Descriptive statistics were applied.

Results: A total of 1,800 advisory consultations were recorded during the study period. The most common categories were post-examination interpretative consultation (35.6%) and clarification of abnormal results (28.3%), followed by pre-examination advice (20.0%) and resolution of inter-laboratory discrepancies (12.2%). Frequently discussed analytes included glucose profiles, lipid parameters, infectious disease serology, and screening tests. Satisfaction data were available for 1,625 consultations, with a mean score of 4.5 ± 0.6; 90% of respondents rated the consultation as satisfactory or very satisfactory.

Conclusions: These findings highlight the growing demand for professional laboratory interpretation in patient-initiated testing environments and support the inclusion of advisory consultations as a measurable quality indicator within laboratory management systems.

Introduction

The expansion of direct-to-consumer (DTC) and patient-initiated laboratory testing has altered the traditional "brain-to-brain" loop of diagnostic medicine [1, 2]. Patients now frequently access laboratory reports without prior clinical interpretation and seek clarification directly from the laboratory [3, 4]. This shift challenges the traditional view of laboratory physicians as background specialists, positioning them as active contributors to patient-centered care. While patient-initiated testing enhances self-empowerment, it raises concerns about unregulated quality, lack of interpretation, and potential risks to patients and public health, including erosion of evidence-based practice standards [5]. An Australian study finds that the majority of DTC tests sold online have low potential clinical utility, often targeting healthy consumers and raising significant concerns about medical overuse, financial harm, and inadequate regulation [6]. The impact of DTC testing may vary depending on context, population, and regulatory environment.

Patient-initiated testing is particularly common in resource-limited settings, where fragmented healthcare, limited access to medical records, and brief consultations lead patients to independently choose tests, thereby increasing the proactive role of laboratory physicians in guiding appropriate test utilization and interpretation. A study in Ghana identifies significant communication gaps between laboratory professionals and patients during pre- and post-sampling interactions, revealing a lack of standardized procedures and patient understanding of laboratory tests despite professionals recognizing the need for better communication [7]. Laboratory physicians, with expertise across pre-analytical, analytical, and post-analytical phases, are uniquely positioned to contextualize results and explain unexpected or discordant findings, particularly when supported by relevant clinical history.

In accordance with ISO 15189:2022, accredited laboratories are required to establish advisory services covering test selection, interpretation of results, and promotion of appropriate test utilization. These responsibilities extend beyond technical validation to include professional consultation. However, limited data exist describing the reasons patients seek consultation with laboratory physicians and how these interactions are perceived. Understanding these interactions is crucial for laboratories to adapt their services, enhance patient care, and navigate the implications of the patient-initiated laboratory testing paradigm. In an author's laboratory, all interpretative consultations are performed by board-certified MD Biochemistry or MD Pathology specialists with competency in clinical interpretation, biological variation, analytical performance characteristics, and harmonization principles. This study aims to categorize the causes and assess the satisfaction associated with patient-driven interactions with laboratory physicians in a clinical laboratory setting. These interactions represent structured advisory services embedded within the laboratory quality management system rather than informal patient encounters.

Materials and Methods

Study Design and Setting

This prospective observational study, evaluating real world demand for laboratory physician consultation, was conducted over a 12-month period (February 2024 to February 2025) at Samyak Diagnostic, a high-volume referral ISO 15189:2022 accredited clinical laboratory in Kathmandu, Nepal. The study was designed to systematically document and analyze real-world, patient-initiated interactions with laboratory physicians occurring during routine service delivery, without altering standard clinical or laboratory workflows or introducing any clinical intervention.

Definition of Patient-Initiated Interaction

A patient-initiated interaction was defined as any direct, unsolicited communication initiated by a patient or accompanying caregiver with a laboratory physician, either before or after testing, for the purpose of test selection, clarification of laboratory results, interpretation of abnormal or unexpected findings, or explanation of inter-laboratory result discrepancies. Interactions initiated by clinicians or internal laboratory quality processes were excluded.

Data Collection Tool and Variables

A structured, purpose designed logbook was developed prior to study initiation and used uniformly by all participating laboratory physicians. For each interaction, the following variables were recorded immediately after consultation:

1. Date of interaction
2. Categorized according to the cause of consultation:

Pre-test consultation

Post-test general consultation Inquiry regarding abnormal results

Inquiry regarding disagreement with reports Inquiry regarding discrepant results from other laboratories

3. Other administrative or procedural queries
4. Primary test(s) involved

To enhance consistency, consultation categories were defined a priori and reinforced through internal orientation sessions. The interactions described represent clinical interpretative consultations rather than routine customer-service helpdesk activities. Consultations involved individualized medical reasoning such as interpretation of discordant or paradoxical results, explanation of biological and analytical variation, assessment of test appropriateness, and advice on need for repeat testing or clinical referral. All critical results during consultation were communicated with the clinicians or referred to the hospital. All abnormal or discordant results discussed during consultations had undergone standard internal quality control verification and, where indicated, repeat analysis prior to release, in accordance with ISO 15189:2022 requirements for ensuring analytical validity.

Patient Satisfaction Assessment

Patient satisfaction was assessed immediately following the interaction using a 5-point Likert scale (1 = very dissatisfied; 5 = very satisfied). As only patients who reached physicians were surveyed, the possibility of positive response bias is acknowledged.

Data Quality Assurance

All participating laboratory physicians received standardized training on documentation procedures before study initiation. Logbook entries were reviewed monthly by the study coordinator

to ensure completeness, internal consistency, and appropriate categorization. Ambiguous entries were resolved by consensus discussion among the laboratory physicians involved.

Ethical Considerations

In accordance with institutional policy, the study was considered exempt from formal ethics committee review. The study is a routine advisory service as per ISO 15189:2022 requirement. Data confidentiality and anonymity were maintained throughout the study. This is in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Statistical Analysis

Data were entered into a secure database and analyzed using descriptive statistics. Frequencies and proportions were calculated for categories of patient interactions. Satisfactionscores were summarized using means and standard deviations. Qualitative feedback was reviewed thematically to identify recurring patterns and illustrative examples and then categorized to respective heading in the log book.

Results

During the study period, 1,800 patient-initiated interactions were recorded. Post-test general consultations were the most frequent (35.6%), followed by clarification of abnormal results (28.3%). Commonly discussed tests included glucose profiles with paradoxical findings, lipid parameters, equivocal infectious disease serology, and prenatal screening results.

Inter-laboratory discrepancies accounted for 12.2% of interactions, most often involving glucose, triglycerides, and thyroid-stimulating hormone. Explicit disagreement with laboratory reports was uncommon (2.8%). The common examples and distribution of causes of interaction is summarized in Table 1.

Table 1: Causes of Patient-Driven Interactions with Laboratory Doctors (n=1800).

Category	Number	Percentage	Commonest Examples
Post-Test general consultation	640	35.6%	Query on highlighted results outside reference interval, Guidance on test frequency and monitoring, Query on additional test and differential diagnosis based on current reports.
Inquiry about abnormal reports	510	28.3%	Paradoxical glucose results, High triglycerides, Equivocal serology, High risk prenatal screening test
Pre-Test consultation	360	20.0%	VDRL, TB Gold, HLA B27, Apo-lipoproteins, FNAC and Biopsy, Reconfirming tests ordered by clinicians for their symptoms, Test selection based on symptoms
Inquiry about different results from another lab	220	12.2%	Triglycerides, TSH, Glucose, Uric acid
Disagreement with reports	50	2.8%	Vitamin B12 levels, HbA1c, Glucose, Uric acid, Blood Tacrolimus level
Others	20	1.1%	Turnaround time, home collection of samples
Total	1800	100%	

Satisfaction data were available for 1,625 interactions. The mean satisfaction score was 4.5 ± 0.6, with 90% of patients rating the consultation as 4 or 5. Lower satisfaction scores were primarily associated with consultations involving conflicting results from different laboratories.

Discussion

Although ISO 15189:2022 mandates advisory services, the volume and pattern of direct patient engagement observed in thisstudy quantify the operational burden and clinical impact of such services, which has been insufficiently documented in the literature. This study demonstrates a substantial and growing demand for direct access to laboratory expertise in routine practice, driven by patient autonomy and the expansion of DTC and patient-initiated testing. This can be considered as operational characterization of advisory services under ISO 15189. Patients most frequently sought clarification for unexpected, abnormal, or discordant results, areas wherelaboratory physicians provide

distinct interpretative value that cannot be replaced by automated reporting alone. The high frequency of pre-test inquiries in our study, for infectious disease serology and screening tests suggests gaps in patient education that could be addressed through improved pre-test information and standardized interpretative comments. The high volume of inquiries on tests like VDRL and TB Gold, often required for employment or travel, underscores a need for better pre-test patient education. Laboratories could mitigate this by developing targeted informational materials, thereby reducing repetitive consultations [8]. The American Sexually Transmitted Diseases Association states that while patient initiated STI testing including VDRL may improve access and reduce stigma, concerns about test reliability, inappropriate test selection, result interpretation, affordability, and integration into surveillance systems necessitate strict quality standards and clear pathways for clinical follow-up as further highlighted by Martin-García and Santi-Rocca (2024) [9,10].

The substantial burden of consultations related to inter-laboratory discrepancies highlights ongoing challenges in result harmonization and analytical comparability, which directly influence patient trust [11]. This was particularly evident for analytes such as glucose, uric acid, and triglycerides, all of which exhibit high intra-individual biological variation. Incorporating statements on measurement uncertainty in laboratory reports may help align clinician and patient expectations with the inherent variability of these tests [12]. Furthermore, several laboratories worldwide provide Reference Change Value (RCV) calculators on their websites to assist clinicians and patients in determining whether observed differences between consecutive results represent statistically significant physiological change rather than random variation [13]. Many laboratories also integrate RCV-based rules within laboratory information systems as part of auto-verification and delta-check algorithms to flag clinically significant changes.

In our study, discrepancies between patient expectations and laboratory results were occasionally observed, predominantly concerning again with paradoxical glucose results (higher fasting glucose versus post-prandial glucose) and unexpected Vitamin B12 levels. Queries regarding elevated fasting glucose represented the vast majority of these disagreements. Providing a clear physiological explanation for this glucose pattern initially presented a considerable difficulty. In response, we published scientific literature addressing this phenomenon and created self-explanatory videos, which were disseminated via social media to educate patients on the causes [14].

International literature presents divergent perspectives on patient-initiated laboratory testing, balancing concerns regarding safety and regulation with recognition of its potential to improve access and self-management. The EFLM DTCT-Taskforce has proposed reframing DTC testing as “Consumer Initiated Testing,” emphasizing that this model circumvents the traditional “brain-to-brain” testing loop and may introduce significant quality risks, particularly when analyses are performed outside medical laboratories [15]. At the same time, the taskforce acknowledges benefits such as improved accessibility, convenience, and consumer choice, especially for underservedpopulations, and recommends dedicated regulation to ensure tests are appropriate for lay use and direct users toward medical consultation rather than autonomous clinical decisions [15]. Ayala-Lopez and Nichols similarly note that consumer demand is driven by convenience, confidentiality, and empowerment, yet caution that absence of clinical context may lead to misinterpretation, privacy concerns, and unnecessary follow-up from false-positive results [16]. Hinzmann argues that patient-initiated testing can support self-management of chronic disease and rapid decision making, the advantages particularly relevant in resource limited settings while advocating that professionals should focus on public education rather than outright rejection of the model [17]. Conversely, Orth et al. highlight substantial risks arising from regulatory gaps, analytical variability, and lack of oversight, asserting that laboratory specialists have an ethical responsibility to inform the public and promote stricter governance to protect diagnostic integrity [18]. While advisory services support test selection and result interpretation, their scope should remain clearly defined, with direct diagnostic and treatment decisions reserved for the treating clinician to ensure appropriate clinical governance and patient safety.

Our findings suggest that laboratory physicians can partially mitigate these risks by guiding appropriate understanding and follow-up. Importantly, the consistently high satisfaction observed indicates that such interactions are not merely informational but perceived as clinically meaningful by patients. This presents a compelling case for laboratories to formalize such consultation services. Early strategies could include:

1. Interpretative reporting: Short, standardized comments for common abnormal or paradoxical results
2. Patient education: Targeted digital or printed materials addressing frequently queried tests
3. Structured consultations: Brief post-test consultations as part of wellness and screening packages

The increasing demand for advisory and interpretative services highlights the need for enhanced training in clinical communication and interpretation skills for laboratory physicians, as well as appropriate staffing models that allocate dedicated time for consultative laboratory practice. These services should be systematically documented within laboratory information systems or structured records, as this not only ensures traceability and continuity of care but also provides an essential medico-legal safeguard within the framework of ISO 15189:2022-accredited laboratory practice.

Patients seeking consultation in our study were heterogeneous, with varying levels of medical literacy that appeared to influence the nature of queries and expectations. While some required basic explanations, others sought detailed interpretation, underscoring the need for adaptable communication by laboratory professionals; future studies should formally assess the impact of health literacy on advisory service utilization. This single center service evaluation may not capture the full range of practices in smaller, non-accredited, or rural laboratories. However, the observed interaction patterns are likely representative of laboratories experiencing increased patient directed testing and reduced clinician mediated interpretation,particularly in similar

resource limited settings. This also highlight the need for future multi-center studies with inferential or comparative designs to validate these observations.

Conclusions

Advisory and interpretative consultations represent a substantial and measurable operational component of ISO 15189:2022-accredited laboratory practice. Patient-initiated consultations with laboratory physicians reflect a growing need for expert interpretation. While these findings are based on a single-center experience and should be interpreted within this context, the quantification of this activity supports its recognition as a potential quality indicator within laboratory management systems.

Acknowledgements

The authors would like to thank Mr. Niroj Karanjit for managing timely patient consultation.

Ethical Approval

The study is a routine advisory service as per ISO 15189:2022 requirement. Data confidentiality and anonymity were maintained throughout the study. This is in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Author Contributions

VP: Conceptualization, Formal analysis, Writing original draft. SP, KG, DP, AS and NN: Data curation, Writing-Review and Editing.

Conflict of interest

None.

Research funding

None.

Data Availability Statements

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

References

1. Karcher DS. Pathologists as clinical consultants: for the patient and1. Karcher DS. Pathologists as clinical consultants: for the patient with the patient. Archives of pathology & laboratory medicine. 2023;147(4):418-24. Doi: arpa.2022-0174-RA
2. Plebani M. The changing scenario of laboratory medicine: the role of lab professionals in the new paradigm. Clin Chem Lab Med. 2023; 61(5):795–801. Doi: 10.1016/j.cca.2008.03.013
3. Pant V. The Evolving Role of Laboratory Doctors: A Shift towards Patient Consultation. JNMA: Journal of the Nepal Medical Association. 2024;62(280):854. Doi: 10.31729/jnma.8838
4. Laposata M, Laposata ME. Patient-specific interpretation of test results: the next frontier. Clin Chem. 2015; 61(4):609–611. Doi: 10.1515/dx-2013-0010
5. Orth M. Direct to consumer laboratory testing (DTCT)–opportunities and concerns. EJIFCC. 2021;32(2):209. PMCID: PMC8343044
6. Shih P, Ding P, Carter SM, Stanaway F, Horvath AR, Langguth D et al. Direct-to-consumer tests advertised online in Australia and their implications for medical overuse: systematic online review and a typology of clinical utility. BMJ open. 2023;13(12):e074205. Doi: 10.1136/bmjopen-2023-074205
7. Adu P. A cross-case analyses of laboratory professionals-patients interaction for patients accessing laboratory services at University of Cape Coast hospital and Ewim Polyclinic in the Cape Coast Metropolis, Ghana. BMC Health Services Research. 2021;21(1):520. Doi: 10.1186/s12913-021-06560-8
8. Geusau A, et al. Biological false-positive tests comprise a high proportion of Venereal Disease Research Laboratory reactions. Int STD AIDS. 2005; 16(11):722-726. Doi: 10.1258/095646205774763207
9. Exten C, Pinto CN, Gaynor AM, Meyerson B, Griner SB, Van Der Pol B. Direct-to-consumer sexually transmitted infection testing services: a position statement from the American Sexually Transmitted Diseases Association. Sexually transmitted diseases. 2021;48(11):e155-9. Doi: 10.1097/OLQ.0000000000001475
10. Martín-García DF, Santi-Rocca J. Direct-to-Consumer Testing: A Game-Changer for STI Control and Public Health? A Critical Review of Advances Since the Onset of the COVID-19 Pandemic. Venereology. 2024;3(4):232-245. Doi: 10.3390/venereology3040018
11. Earley MC, et al. Practices and perceived value of proficiency testing in clinical laboratories. J Appl Lab Med. 2017; 1(4):415-520. Doi: 10.1373/jalm.2016.021469
12. Plebani M, et al. What information on measurement uncertainty should be communicated to clinicians, and how? Clin Biochem. 2018; 57:18-22. Doi: 10.1016/j.clinbiochem.2018.01.017
13. https://www.furst.no/analyse-og-klinikk/alfabetisk-analyseliste/hba1c (Accessed on 16/01/2026)
14. Pant V, Gautam K, Pradhan S. Postprandial Blood Glucose can be less than Fasting Blood Glucose and this is not a Laboratory Error. JNMA: Journal of the Nepal Medical Association. https://www.furst.no/analyse-og-klinikk/alfabetisk-analyseliste/hba1c2019;57(215):67. Doi: 10.31729/jnma.3985
15. Shih P, Sandberg S, Balla J, Basok BI, Brady JJ, Croal B et al. Direct-to-consumer testing as consumer initiated testing: compromises to the testing process and opportunities for quality improvement: An opinion paper from the EFLM DTCT-Taskforce. Clinical Chemistry and Laboratory Medicine (CCLM). 2025; 63(2):262-269. Doi: 10.1515/

cclm-2024-0876

16. Ayala-Lopez N, Nichols JH. Benefits and risks of direct-to-consumer testing. Archives of Pathology & Laboratory Medicine. 2020; 144(10):1193-1198. Doi: 10.5858/arpa.2020-0078-RA
17. Hinzmann R. Direct-to-consumer testing–benefits for consumers, people with disease and public health. Clinical Chemistry and Laboratory Medicine (CCLM). 2023; 61(4):703-708. Doi: 10.1515/cclm-2023-0007
18. Orth M, Sandberg S, Shih P. Direct-to-Consumer Testing: Benefits and Concerns of Commercially Accessed Laboratory Tests. Clinical Chemistry. 2025; 71(6):652-663. Doi: 10.1093/clinchem/hvaf004

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research article

Targeting Sample Reception Errors: A Failure Reporting Analysis and Corrective Action System (FRACAS)-Based Laboratory Quality Improvement Study

Vijetha Shenoy Belle[1], Merin K Thomas[2], Nihaal Maripini[3], Saritha Kamath U[2*]

[1]Department of Biochemistry, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, India
[2]Department of Medical Laboratory Technology, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India
[3]Department of Biochemistry, Great Eastern Medical School and Hospital, Srikakulam, India

Article Info

Corresponding Author:
Saritha Kamath U
Professor, Department of Medical Laboratory Technology, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India
Phone: +91 9844344459
E-mail: saritha.kamath@manipal.edu

Keywords

FRACAS, Pre-examination error, Risk Priority Number, Quality improvement, Risk management, Quality indicators, SDG-4

Abstract

Introduction: Approximately 70% of clinical decisions depend on laboratory test results, underscoring the critical importance of examination accuracy. Among the phases of the total testing process, the pre-examination phase accounts for the highest proportion of errors; therefore, interventions aimed at reducing these errors can significantly enhance overall laboratory quality. This study aims to evaluate the effectiveness of risk assessment and corrective strategies through the implementation of the Failure Reporting, Analysis, and Corrective Action System (FRACAS) by comparing Risk Priority Number (RPN) scores before and after implementing corrective strategies.

Methodology: This retrospective observational quality-improvement study obtained approval from the institutional ethics committee. It analyzed two years of data from a clinical biochemistry laboratory. A total of 573 samples were rejected in a NABL-accredited, ISO 15189-compliant tertiary-care hospital in South India. Each nonconformance (NC) was categorized, and its frequency, percentage distribution, and RPN were calculated for both study periods. The effectiveness of risk assessment and corrective strategies was evaluated by comparing pre- and post-corrective RPN values and percentage reductions in NCs.

Results: Data from 573 rejected samples (216 pre-corrective in 2021; 357 post-correctives in 2022) were analysed. Mean RPN decreased from 18.0 ± 16.1 to 10.9 ± 15.3 (mean difference 6.0, 95% CI 2.0–24.0, $p = 0.041$); paired Bayesian t-test showed $BF_{10} = 1.15$, evidence for effectiveness. Four NCs were eliminated; however, three new NCs emerged post-STAT lab launch.

Conclusion: FRACAS-guided interventions reduce pre-examination NCs and lower RPNs for high-risk processes, supporting their use for pre-examination quality improvement.

Introduction

Accurate clinical laboratory testing is fundamental to effective patient diagnosis, treatment, and monitoring. In healthcare settings, approximately 70% of medical decisions are based on laboratory test results, highlighting the critical importance of timely and accurate reporting. The total testing process (TTP) in clinical laboratory medicine comprises three phases: pre-examination, examination, and post-examination [1]. Optimal management of each phase is essential to ensure the generation of precise and reliable laboratory results [2].

Advances in laboratory automation and technology have substantially reduced errors in the examination and post-examination phases. In contrast, the pre-examination phase, largely influenced by human factors and frequently occurring outside the laboratory, remains the predominant source of errors within the TTP and therefore requires focused improvement. The pre-examination phase encompasses all processes from the clinician's test request to the initiation of sample analysis [2]. Errors may occur at any stage of the TTP; however, pre-examination errors account for the majority, representing approximately 46–68.2% of total laboratory errors, compared with 7–13% for examination errors and 18.5–47% for post-examination errors [3]. Consequently, laboratory leadership bears the responsibility of maintaining risks at clinically acceptable levels. Quality improvement initiatives include systematic identification, evaluation, corrections, and continuous monitoring of errors [4]. Within the pre-examination phase, errors most commonly arise during patient preparation and specimen collection, underscoring the need for a skilled phlebotomy workforce and ongoing training of personnel involved in pre-examination activities [4,5]. Quality indicators (QIs) play a pivotal role in continuous quality improvement by enabling monitoring and reduction of errors across all phases of the TTP [6].

Risk identification and analysis can be effectively performed using the Failure Reporting, Analysis, And Corrective Action System (FRACAS), which categorizes nonconformances (NCs) based on their Risk Priority Number (RPN). The RPN is calculated as the product of occurrence, severity, and detection scores and facilitates risk prioritization and stratification [3]. Failure Reporting Analysis and Corrective Action System is widely recognized as a robust and systematic approach to risk management. Within FRACAS, every reported failure initiate a thorough investigation. After the root cause is determined, corrective actions are implemented to resolve the immediate issue. Preventive measures are subsequently introduced to reduce the likelihood of recurrence, often through updates to procedures, enhanced training, or modifications to equipment configuration [7]. Ensuring patient safety is a primary objective of healthcare systems, and initiatives aimed at reducing diagnostic errors directly contribute to improved patient outcomes [8]. Data on FRACAS-driven risk assessment specifically targeting sample reception areas of the pre-examination NCs in Indian tertiary-care biochemistry laboratories are limited. This study aims to demonstrate the application of risk management principles and targeted corrective strategies in clinical laboratories to prevent or reduce errors, with particular emphasison the sample reception area of the pre-examination phase, thereby enhancing the overall quality of patient care.

Materials and methods

This retrospective record-based study analysed documented pre-examination NCs focusing on the sample reception area of the laboratory, from January 2021 (pre-corrective baseline) to December 2022 (post-corrective) in the Department of Biochemistry NABL-accredited, ISO 15189-compliant tertiary-care hospital in South India. The study was conducted after obtaining approval from the "Institutional Research Committee (IRC)" and "Institutional Ethics Committee" (IEC 360/2023). The study included documented data on NCs from the laboratory's sample reception area from January 2021 to December 2022 and excluded data on examination and post-examination non-conformances in the clinical biochemistry laboratory.

Data collection

A complete enumeration of data from January 2021 to December 2022 was taken; each NC event was counted once, irrespective of the number of analytes ordered on that specimen. Data were compiled and categorised to quantify the frequency of each pre-examination NC reported in the lab. The total number of tests processed increased from 2,619,345 in 2021 to 3,268,618 in 2022, indicating a rise in workload during the latter year. Similarly, the number of sample collection tubes received in the laboratory increased from 187,096 in 2021 to 233,472 in 2022. The total number of pre-examinations NCs recorded was 216 in 2021 and 356 in 2022. When calculated relative to the total number of samples received, the rejection rates were 0.115% in 2021 and 0.152% in 2022, demonstrating a slight increase in the proportion of pre-examination errors in 2022. Data between January 2021 and December 2021 were considered as the baseline (pre-correction period) for comparative purposes. Specimens included in the study comprised those obtained from outpatient departments and all the wards of the hospital.

Design

The reported NCs were categorised into 18 distinct pre-examination NC types based on their frequency of occurrence for both study years. Severity rate (SR), occurrence rate (OR), and detection rate (DR) were assigned (1-5 scale) by a multidisciplinary team (laboratory director, quality manager, senior biochemist) using ISO 22367:2020 criteria via consensus. Risk Priority Numbers were calculated as RPN = SR × OR × DR, and RPN bands were defined: ≥40 high risk, 10-39 moderate, <10 low.

Training: Training is a continuous process in laboratory practice. As part of the corrective strategy implemented in this study, targeted training programs were conducted for all pre-examination NCs identified in 2021. The identified errors were systematically communicated to the respective wards, phlebotomy units, or reception areas. Section in-charges were informed and instructed to provide appropriate training addressing the specific errors and to maintain records of the corrective actions undertaken. Training was conducted during the monthly review meeting of the laboratory for approximately 30minutes, involving an

average of 10 technicians. The session focused on discussing errors identified in the sample reception area, analyzing their root causes, and outlining mitigation strategies to minimize such errors in the future. For the NC of "sample mislabelled," the root cause analysis identified factors such as skipping double-check steps due to high workload, inconsistent practices among technicians, and multitasking during peak hours. The mitigation plan included implementing mandatory double-check procedures, displaying standardized protocols in the sample reception area, ensuring uniform adherence to procedures by all technicians, and optimizing workload distribution during peak periods. Similarly,standardized procedures were prepared for all the other NCs, and training was conducted to follow them.

Improvements were quantified by calculating percentage reductions using the 2021 RPN values as the baseline. The RPN analysis provided insight into prioritizing areas requiring intervention to eliminate or minimize pre-examination errors. Comparison of pre- and post-corrective RPN values enabled evaluation of the effectiveness of the implemented risk assessment and corrective action plans.

Each NC was assessed in terms of severity, occurrence, and detection. Scores ranging from 1 to 5 were assigned for each parameter based on predefined criteria, as detailed in Table 1 [9].

Table 1: Severity rate, Occurrence rate and Detection rate scoring details.

Level	Term	Description
Severity Rate Score		
5	Critical	Significant adverse clinical outcome: Life-threatening injury/death.
4	Serious	Moderate adverse clinical outcome: Irreversible bodily damage.
3	Significant	Minor adverse clinical outcome: Temporary physical damage or disability can be remedied with medical treatment.
2	Marginal	No negative clinical outcome: Transient physical damage or impairment, can be reversed without medical intervention.
1	Insignificant	No adverse clinical outcome: Unchanged patient management.
Occurrence Rate Score		
5	Frequent	Happens each day. Very likely to happen soon, with consistent exposure.
4	Reasonably Likely	Each week. Probable to happen soon, with routine exposure expected.
3	Occasional	Each month: Possible to happen at some point in the predictable future, occasional exposure is possible.
2	Remote Each year	Each year; Unlikely to happen in the foreseen future
1	Unlikely	Less than once a year. Will happen only in rare or extraordinary situations.
Detection Rate Score		
5	Very low	The detection rate is minimal, indicating that the current controls are unlikely to identify the presence of the failure mode.
4	Low	There's a low likelihood of identifying the presence of the failure mode with the current controls.
3	Moderate	There is a possibility of identifying the presence of the failure mode with the current controls.
2	High	There is a strong likelihood of identifying the failure mode with the current controls.
1	Very high	It is almost certain to identify the failure mode with the current controls.

Statistical analysis
The percentage and RPN for each NC were calculated and described in Table 2. Pre- and post-corrective RPN for 15 common NCs (n=15 paired observations) were compared using paired t-tests; normality of RPN differences was checked via Shapiro-Wilk test. The effect of the study was assessed using a Bayesian paired sample t-test with default Cauchy prior. Jamovisoftware (Version 2.3.28) was used to do the statistical analysis. Statistical significance was set at $p < 0.05$, and $BF_{10} > 1$ indicated evidence against the null hypothesis.

Results
In the present study, pre-corrective data (2021) and post-corrective data (2022) were considered.

Table 2: Preexamination NC of 2021 and 2022 in number and percentage.

Non- conformances	2021 Numbers (%)	2022 Numbers (%)
Samples lost/not received	70 (32.41%)	115 (32.21%)
Improperly labelled samples	49 (22.69%)	36 (10.08%)
Sample without requisition slip	19 (8.80%)	11 (3.08%)
Sample received in wrong preservative/ wrong vacutainer	18 (8.33%)	32 (8.96%)
Inappropriate sample	16 (7.41%)	20 (5.60%)
Test not billed	12 (5.56%)	3 (0.84%)
Quantity not sufficient except for NICU	11 (5.09%)	30 (8.40%)
Wrong test billed	6 (2.78%)	3 (0.84%)
Test missed	5 (2.31%)	7 (1.96%)
Improper sample transport	4 (1.85%)	4 (1.12%)
Test request form without label	2 (0.93%)	0 (0.00%)
Delay in sample reception	1 (0.46%)	0 (0.00%)
Unordered repeat sample	1 (0.46%)	0 (0.00%)
Bill cancelled	1 (0.46%)	4(1.12%)
Not an in-house test	1 (0.46%)	0 (0.00%)
Barcode not stuck properly	0 (0.00%)	1 (0.28%)
Clotted sample	0 (0.00%)	79 (22.13%)
TRF duplication	0 (0.00%)	12 (3.36%)
Total	216 (100%)	357 (100%)

As a part of risk assessment, RPN was calculated. RPN provides an insight into the clinical impact of NC in patient care. Pre- and post-corrective strategy RPN is mentioned in Table 3.

Table 3: Pre- and post-corrective strategy RNP scores.

2021 and 2022 non-c onformanc es	OR	DR	SR	RPN 2021	Corrective Strategies	OR	DR	SR	RPN 2022
Sample received in the wrong p reservative/ improper(w rong) vacutainer	3	3	5	45	Targeted Training and Compet ency Assessment	4	3	5	60
Improperly labelled samples	3	3	5	45	Reinforce patient iden tification protocols (two identifiers) at collection Mandatory bedside/bar code labeling im mediately after collection Periodic training	3	2	2	12
Samples lost/ not received	3	3	5	45	Clear handover re sponsibility between the collection and transport staff	3	1	1	3
Test missed	2	3	5	30	Checking the LIS system for pending or unprocesse d tests at the end of each duty shift.	2	1	5	10
Wrong test billed	2	4	3	24	LIS validation checks during billing and staff training	2	4	3	24

2021 and 2022 non-c onformanc es	OR	DR	SR	RPN 2021	Corrective Strategies	OR	DR	SR	RPN 2022
Test not billed	3	3	2	18	Training of a billing clerk	2	3	2	12
Improper sample transport	2	2	4	16	Training -Educate transport personnel on sample stability req uirements	2	2	4	16
Quantity not sufficient except for NICU	3	1	4	12	• Display minimum volume req uirements at collection areas • Phlebotomy training on correct draw volumes	3	1	3	9
Inappropria te sample	3	1	3	9	Display of Sample type and container re quirements LIS prompts for correct sample selection Training	3	1	3	9
Delay in sample reception	2	1	3	6	Monitor delays as a quality indicator	1	1	1	1
Sample without requisition slip	3	1	2	6	• Enforce mandatory requisition at sample acceptance • LIS-based electronic test ordering	3	1	2	6
Test request form without label	2	1	3	6	Training Display of the procedure in the workstation	1	1	1	1

2021 and 2022 non-c onformanc es	OR	DR	SR	RPN 2021	Corrective Strategies	OR	DR	SR	RPN 2022
Unordered repeat sample	2	2	1	4	Staff education on test ordering protocols	1	1	1	1
Bill cancelled	2	1	1	2	• Clear com munication with patient s/wards before processing	2	1	1	2
Not an in-house test	2	1	1	2	Staff training at reception	1	1	1	1
The barcode is not stuck properly	1	1	1	1	Mandatory visual check before sample dispatch • Training on correct barcode placement	2	2	3	12
Sample clotted	1	1	1	1	Training on correct blood collection technique and mixing	4	1	4	16
TRF duplication	1	1	1	1	Educate staff on avoiding repeat form generation	2	1	1	2

Table 4: List of Non-conformances disappeared in 2022 after corrective strategies.

NCs disappeared in 2022	RPN 2021	RPN 2022
Test request form without label	6	1
Delay in sample reception	6	1
Unordered repeat sample	4	1
Not an in-house test	2	1

Table 5: Comparison between pre- and post-RPN corrective strategies using paired t-test.

RPN	Mean ± SD	Mean difference	SE difference	95% Confidence Interval	95% Confidence Interval	Effect Size	P
Before Corrective Strategy (in 2021)	18±16.1	6	3.69	2	24	0.745	0.041
After Corrective Strategy (in 2022)	10.9±15.3	6	3.69	2	24	0.745	0.041

The clinical relevance of the study was assessed using "Bayesian paired t-test", detailed in Table 6.

Table 6: Showing the confidence interval and Bayser factor.

RPN	Mean ± SD	95% Confidence Interval	95% Confidence Interval	BF	Error %
Before corrective Strategy (in 2021)	18±16.1	9.09	26.9	1.15	0.0242
After corrective Strategy (in 2022)	10.9±15.3	2.42	19.3	1.15	0.0242

The Bayes factor (BF) provides evidence that supports one hypothesis over another [10]. In this study, the alternative hypothesis (H1) states that there is a significant effectiveness noted in the preexamination phase following the implementation of the corrective plan, while the null hypothesis (H0) claims there is no significant effectiveness. BF_{10}=1.15 claims H1 has evidence over H0. BF_{10} and a lower error % of 0.0242 support the study to be clinically significant.

Discussion

The total number of sample collection tubes received in the laboratory increased from 187,096 in 2021 to 233,472 in 2022, accompanied by a rise in pre-examination NCs from 216 to 356. When expressed relative to the total sample volume, the rejection rate showed a slight increase from 0.115% in 2021 to 0.152% in 2022. These findings provide an overview of the trend in pre-examination NCs across the study period.

Failure Reporting Analysis and Corrective Action System-guided interventions were associated with the reduction or elimination of several high-risk pre-examination NCs for e.g., samples lost/not received: High risk RPN to low risk (RPN 45→3); improperly labelled: high risk RPN to moderate risk (RPN 45→12), with overall mean RPN falling significantly (18.0±16.1 to 10.9±15.3, p=0.041). Of 15 common NCs, 9 showed decreased RPN, 4 remained unchanged, and 2 increased. Four NCs disappeared entirely, though three new NCs emerged coinciding with STAT lab launch, highlighting that workflow changes can introduce new error modes. These findings demonstrate that FRACAS-driven risk assessment combined with targeted corrective strategies led to measurable improvement in several pre-examination processes, though persistent and worsening NCs indicate that training alone may be insufficient for complex systemic issues.

The pre-examination phase remains the most error-prone component of the TTP, largely due to its dependence on human factors and processes occurring outside the examination environment [1]. Studies have shown that preexamination errors outnumber examination or post-examination errors by 4 to 6 times in many countries, prompting initiatives for error reduction [12,13,14,15].

Among the identified NCs, "samples lost/not received" represented the most significant risk before corrective actions, with a high RPN of 45 in 2021. This finding contrasts with lower rates reported in earlier studies, where the contribution of this error ranged from 1.5% to 13.21% [2,16]. Following the implementation of defined handover responsibilities and improved communication between collection and transport staff, the RPN decreased markedly to low risk (RPN 3) in 2022, indicating a reduction in both occurrence and detectability risk. This underscores the importance of accountability and traceability in sample handling workflows. "Improperly labelled samples" also constituted a major pre-examination error in 2021, with an RPN of 45 indicating highrisk. After reinforcing patient identification

protocols, implementing mandatory bedside labelling, and providing periodic training, the RPN decreased significantly to moderate risk (RPN 12) in 2022. Although the initial rate was higher than those reported in previous studies [2,11,16], the post-corrective strategy reduction aligns with findings from other investigations that demonstrated improved labelling accuracy following targeted staff education [17,18]. This highlights training as an effective corrective strategy for human-factor–related errors.

Errors related to “wrong preservative or improper vacutainer” remained a persistent challenge. Despite a corrective strategy through targeted training and competency assessment, the RPN increased from 45 in 2021 to 60 in 2022, both are in high-risk category due to a rise in occurrences as well as increased samples received in 2022. Similar trends of increased frequency following interventions have been reported in other studies [17], suggesting that this NC may require additional strategies beyond training, such as standardized collection kits, stronger LIS prompts, or visual aids at collection sites. This finding emphasizes that training alone may not be sufficient for certain high-risk errors. Another reason for the increased number of occurrences is improved documentation of the NCs.

Billing-related NCs demonstrated mixed outcomes. “Wrong test billed” maintained a moderate risk RPN of 24 in both years despite a reduction in frequency, reflecting persistent challenges in detection and severity. The “test not billed” issue showed a decrease in RPN from 18 to 12 after targeted training of billing personnel, reflecting improved process reliability. However, the RPN score remains within a moderate risk range. These results indicate that mitigating billing-related risks requires a combination of system-level controls and ongoing staff oversight. Errors related to sample transport and reception showed improvements. “Improper sample transport” exhibited no change in RPN despite a reduction in frequency, consistent with reports that transportation errors can significantly compromise sample integrity if not tightly controlled [19]. In contrast, “delay in sample reception” and “test request form without label” were eliminated following the corrective strategy, with RPNs reduced to one, demonstrating the effectiveness of monitoring these parameters as quality indicators and enforcing standardized procedures.

Certain NCs emerged only after workflow changes. “Clotted samples” appeared in 2022 following the introduction of a short turnaround time (STAT) laboratory service, contributing moderate risk RPN of 16. This finding is comparable to previous studies reporting high rates of clotted samples [11]. Inadequate mixing of anticoagulant tubes after collection is a likely contributing factor [20], highlighting the need for focused phlebotomy training when new services or workflows are introduced.

Overall, this study demonstrates that incorrect techniques, workflow gaps, and communication failures are key contributors to pre-examination errors. Regular monitoring of NCs and RPNs enables laboratories to prioritize risks effectively and evaluate the impact of corrective actions. While training proved effective for many NCs, some high-risk errors persisted or increased, indicating the need for multifaceted interventions that combineeducation, system automation, standardized procedures, and continuous performance monitoring.

Continuous staff training strengthened laboratory–clinical communication, implementation of quality indicators, and routine review of RPN trends are essential to sustaining improvements in the pre-examination phase. Adoption of a structured risk management approach, such as FRACAS, supports proactive error prevention, enhances laboratory efficiency, and ultimately contributes to improved patient safety and quality of care [21,22].

Conclusion

Implementation of FRACAS-driven risk assessment and targeted corrective strategies significantly reduced pre-examination non-conformances and associated risks in the clinical laboratory. Continuous monitoring of RPN enabled effective prioritization of high-risk processes and evaluation of corrective actions. Sustained quality improvement in the pre-examination phase requires ongoing staff training, system-level controls, and proactive risk management to enhance patient safety and laboratory performance.

Limitations

In the study, only the biochemistry laboratory's NCs (573) were analysed. Including all NCs from other laboratories in the hospital could offer a more comprehensive view. The study's scope is restricted to data comprising only two years. Extending the duration of the study could yield further valuable insights.

Declaration of Conflict of interests

We all authors declare no conflict of interests.

Ethics statement

The study was conducted after obtaining approval from the Institutional Ethics Committee (IEC 360/2023). This study used only anonymized laboratory quality-related data and did not involve any direct patient participation or identifiable patient information. The study adhered to the principles of the Declaration of Helsinki.

Acknowledgement

AI tool ChatGPT and Perplexity are used for paraphrasing the texts in the manuscript.

Author's Disclosures

Nothing to disclose.

Credit Author Statements

VSB: Concept, project administration, supervision, analysis, manuscript editing and review. MKT: Concept, data collection and analysis, manuscript writing and review. NM: Concept, data analysis, manuscript editing and review. SKU: Concept, supervision, manuscript editing and review.

Funding

No funding obtained.

Data Availability

Data is available with the corresponding author and can be shared upon a reasonable request from the readers.

References

1. Sonmez C, Y·ld·z U, Akkaya N, Taneli F. Preanalytical Phase Errors: Experience of a Central Laboratory. Cureus. 2020;12(3):e7335. doi: 10.7759/cureus.7335. PMID: 32313776; PMCID: PMC7164707.
2. Mehndiratta M, Pasha EH, Chandra N, Almeida EA. Quality Indicators for Evaluating Errors in the Preanalytical Phase. J Lab Physicians. 2021;13(2):169-174. doi: 10.1055/s-0041-1729473. Epub 2021 May 26. PMID: 34483565; PMCID: PMC8409113.
3. Eliza DR, Minodora D. Risk management in clinical laboratory: from theory to practice. Acta Marisiensis-Seria Medica. 2015;61(4):372-377. DOI:10.1515/amma-2015-0086
4. Sinici Lay I, Pınar A, Akbıyık F. Classification of reasons for rejection of biological specimens based on pre-preanalytical processes to identify quality indicators at a university hospital clinical laboratory in Turkey. Clin Biochem. 2014;47(12):1002-1005. doi: 10.1016/j.clinbiochem.2014.04.024. Epub 2014 May 2. PMID: 24794787.
5. Lima-Oliveira G, Lippi G, Salvagno GL, Picheth G, Guidi GC. Laboratory Diagnostics and Quality of Blood Collection. J Med Biochem. 2015;34(3):288-294. doi: 10.2478/jomb-2014-0043. Epub 2015 Jul 14. PMID: 28356839; PMCID: PMC4922344.
6. Sciacovelli L, Plebani M. The IFCC Working Group on laboratory errors and patient safety. Clin Chim Acta. 2009;404(1):79-85. doi: 10.1016/j.cca.2009.03.025. Epub 2009 Mar 26. PMID: 19328194.
7. Alotaibi AME, Hamdi MH, Alomi AM, Masmali AY, Alzahrani ASA, Alshaikhi MAI et al. Assessment of risk management practices in clinical laboratories and their impact on patient safety. J Angiother. 2023;7(1):1–10. https://doi.org/10.25163/angiotherapy.8110327
8. Lippi G, Simundic AM, Mattiuzzi C. Overview on patient safety in healthcare and laboratory diagnostics. Biochemia Medica. 2010;20(2):131-142. https://doi.org/10.11613/BM.2010.015
9. International Organization for Standardization. ISO 22367:2020: Medical laboratories - Application of the risk management to medical laboratories. Geneva, Switzerland: International Organization for Standardization; 2020. https://www.iso.org/standard/71254.html
10. Wei Z, Yang A, Rocha L, Miranda MF, Nathoo FS. A Review of Bayesian Hypothesis Testing and Its Practical Implementations. Entropy (Basel). 2022;24(2):161. doi: 10.3390/e24020161. PMID: 35205456; PMCID: PMC8871131.
11. Dundar C, Bahadir O. Preanalytical Errors in Clinical Biochemistry Laboratory and Relationship With Hospital Departments and Staff: A Record-Based Study. J Patient Saf. 2023;19(4):239-242. doi: 10.1097/PTS.0000000000001115. Epub 2023 Feb 27. PMID: 36849433.
12. Erdal EP, Mitra D, Khangulov VS, Church S, Plokhoy E. The economic impact of poor sample quality in clinical chemistry laboratories: results from a global survey. Ann Clin Biochem. 2017;54(2):230-239. doi: 10.1177/0004563216651647. Epub 2016 Sep 28. PMID: 27166314.
13. van Balveren JA, Verboeket-van de Venne WPHG, Erdem-Eraslan L, de Graaf AJ, Loot AE, Musson REA et al. Dutch Society for Clinical Chemistry and Laboratory Medicine, task group 'SMILE': Signalling Medication Interactions and Laboratory test Expert system. Diagnostic error as a result of drug-laboratory test interactions. Diagnosis (Berl). 2019;6(1):69-71. doi: 10.1515/dx-2018-0098. PMID: 30753158.
14. Meddeb R, Pisareva E, Thierry AR. Guidelines for the Preanalytical Conditions for Analyzing Circulating Cell-Free DNA. Clin Chem. 2019;65(5):623-633. doi: 10.1373/clinchem.2018.298323. Epub 2019 Feb 21. PMID: 30792266.
15. Lippi G, von Meyer A, Cadamuro J, Simundic AM. Blood sample quality. Diagnosis (Berl). 2019;6(1):25-31. doi: 10.1515/dx-2018-0018. PMID: 29794250.
16. Jafari E, Malekpour Afshar R, Aminzade R. Rates and Reasons of Laboratory Sample Rejection due to Pre-analytical Errors in Clinical Settings. Arch Iran Med. 2022;25(3):166-170. doi: 10.34172/aim.2022.28. PMID: 35429958.
17. Arslan FD, Karakoyun I, Basok BI, Aksit MZ, Celik E, Dogan K et al. The Effects of Education and Training Given to Phlebotomists for Reducing Preanalytical Errors. J Med Biochem. 2018;37(2):172-180. doi: 10.1515/jomb-2017-0045. PMID: 30581354; PMCID: PMC6294099.
18. Al-Ghaithi H, Pathare A, Al-Mamari S, Villacrucis R, Fawaz N, Alkindi S. Impact of Educational Activities in Reducing Pre-Analytical Laboratory Errors: A quality initiative. Sultan Qaboos Univ Med J. 2017;17(3):e309-e313. doi: 10.18295/squmj.2017.17.03.008. Epub 2017 Oct 10. PMID: 29062553; PMCID: PMC5642360.
19. Cakirca G. The Evaluation of Error Types and Turnaround Time of Preanalytical Phase in Biochemistry and Hematology Laboratories. Iran J Pathol. 2018;13(2):173-178. Epub 2018 Jul 17. PMID: 30697287; PMCID: PMC6339495. https://pmc.ncbi.nlm.nih.gov/articles/PMC6339495/
20. Arul P, Pushparaj M, Pandian K, Chennimalai L, Rajendran K, Selvaraj E et al. Prevalence and types of preanalytical error in hematology laboratory of a tertiary care hospital in South India. J Lab Physicians. 2018;10(2):237-240. doi: 10.4103/JLP.JLP_98_17. PMID: 29692594; PMCID: PMC5896195.
21. Lillo R, Salinas M, Lopez-Garrigos M, Naranjo-Santana Y, Gutiérrez M, Marín MD et al. Reducing preanalytical laboratory sample errors through educational and technological interventions. Clin Lab. 2012;58(9-10):911-917. PMID: 23163106. https://pubmed.ncbi.nlm.nih.gov/23163106/

22. Lippi G, Baird GS, Banfi G, Bölenius K, Cadamuro J, Church S et al. Improving quality in the preanalytical phase through innovation, on behalf of the European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working Group for Preanalytical Phase (WG-PRE). Clin Chem Lab Med. 2017;55(4):489-500. doi: 10.1515/cclm-2017-0107. PMID: 28231060.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

Integration of Patient-Based Real-Time Quality Control with Conventional Internal Quality Control: Improved Error Detection and Cost-Benefit Analysis in a Clinical Laboratory

Mudasir Bashir Dandroo[1] ,Devanatha Desikan V[1] ,Ramesh Ramasamy[1*]

[1]*Department of Biochemistry, Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER), Puducherry, India*

Article Info

[*]***Corresponding Author:***

Ramesh Ramasamy
Professor and Head, Department of Biochemistry Jawaharlal Institute of Postgraduate Medical Education and Research (JIPMER)
Dhanvantari Nagar, Puducherry, 605006, India

E-mail: rameshrdr30@gmail.com

Phone: 9488365657

Keywords

Patient-Based Real-Time Quality Control (PBRTQC), Internal Quality Control (IQC), Moving Average (MA), Error Detection, Sigma Metrics, Bias Detection, Truncation Limits, Average number of patients data required for error detection (ANPed), Clinical Biochemistry, Laboratory Quality Management, Cost-Benefit Analysis, Risk Management, Clinical Laboratory Improvement Amendments (CLIA)

Abstract

Background: Internal Quality control (IQC) is essential as it ensures analytical accuracy. However, IQC may have certain limitations, for example, it may miss intermittent or matrix-related errors. Patient-Based Real-Time Quality Control (PBRTQC) can overcome this limitation by continuously monitoring the patient data and also aids in early detection of shifts. Hence, integration of PBRTQC with IQC enhances the error detection and the cost efficiency in clinical laboratories.

Aim: To study the effect of integrating Patient-Based Real-time Quality Control with Conventional Quality Control for improved detection of errors, cost benefit and risk-based cost benefits.

Methods: Four parameters, creatinine, urea, Aspartate Transaminase (AST), and Thyroid Stimulating Hormone (TSH), were selected based on low sigma value and high monthly sample load. Historical patient data was collected from the laboratory information system. Data was divided into training and validation sets. Truncation limits were calculated using the interquartile range method. Simple Moving Average (SMA) was calculated for various block sizes. Control limits were set using the maximum/minimum of the moving average or percentiles. A bias simulation study was performed by introducing bias from -50% to +50% to plot bias detection curves and determine the optimum block size and ANPed. Optimized PBRTQC parameters were validated. The optimized algorithm was applied to daily patient data and integrated with IQC. Moving average alarm rate and percentage increase in error detection were calculated. Potential cost-benefit and risk-based benefits were determined.

Results: optimum block size and (ANPed) at TEa was 50 for creatinine (161), 25 for urea (207), 50 for AST (22), and 25 for TSH (68). The MA alarm percentage remained below 1% for all analytes. Integration improved error detection by 28.57% for creatinine, 25% for urea, and 20% for TSH. Net cost benefit demonstrated total savings of 426.92 through QC reductions (264.56) and avoided reanalysis costs (162.36).

Conclusion: PBRTQC integration with IQC significantly improves error detection for low sigma performance assays, is cost-effective and provides a strong risk-based cost-benefit advantage.

Introduction

Quality control (QC) represents a fundamental component of the laboratory quality management system and is essential for maintaining analytical accuracy and reliability. Within this framework, Internal Quality Control (IQC) serves as a core element that provides daily, routine monitoring of analytical performance, ensuring that patient results are valid and reliable before release [1]. Despite its importance, IQC alone does not always guarantee consistent analytical reliability. Several well-recognised limitations exist. First, because IQC is performed at intervals, it cannot detect analytical errors that occur between two QC events. Second, matrix effects hidden interferences within control materials can cause non-commutability, meaning that control samples may not accurately reflect patient specimens. Third, some assays lack stable, commutable control materials, limiting their monitoring effectiveness. Finally, methods with sigma metrics below 4 are inherently more prone to analytical error, and conventional IQC strategies may not provide sufficient assurance of quality in such cases [2,3].

To address these gaps, many laboratories are now incorporating Patient-Based Real-Time Quality Control (PBRTQC) systems. PBRTQC is a modern adaptation of the classical “Average of Normals” (AON) concept introduced by Hoffman and Waid. By applying advanced statistical algorithms to aggregated patient data, PBRTQC enables continuous, real-time monitoring of analytical performance. This approach not only complements traditional IQC but also provides coverage across the pre-analytical, analytical, and post-analytical phases, making it a more comprehensive quality safeguard [4–6].

PBRTQC is particularly useful for analytes that exhibit low biological variation (CVg) compared with analytical variation (CVa), which is typical of tests with lower sigma performance [7,8]. However, successful implementation requires carefulunderstanding of the laboratory’s patient population, including age, sex distribution, and the types of clinical services provided. When tuned to these characteristics, PBRTQC can detect even subtle shifts in analytical performance that may otherwise go unnoticed.

The statistical foundation of PBRTQC includes several critical elements: algorithms for calculating continuous or moving patient means, determination of block size, setting truncation limits to exclude physiological outliers, and defining control limits for identifying and validating potential analytical issues. When combined, these features establish PBRTQC as a robust, data-driven tool for continuous quality monitoring in modern clinical laboratories.

This study aimed to implement a PBRTQC protocol using the moving average algorithm for four key analytes and integrate it with existing IQC procedures, and evaluate its performance in improving error detection and cost effectiveness.

Materials and methods

Study Design and Setting

This study was conducted in the Department of Biochemistry, JIPMER, Puducherry. Institutional ethical approval was obtained (JIP/IEC-OS/2023/053) with a waiver of consent.

Analyte Selection and Data Collection

Four analytes creatinine, urea, AST, and TSH - were selected based on a sigma value <4, AST (sigma 4.39) was included as a comparator to help determine whether the benefits of PBRTQC integration are specific to low-sigma assays or represent a more general effect that might also benefit higher performing assays and high monthly sample volume to ensure robust data analysis. Five months of historical patient data were collected from the Laboratory Information System (LIS). Data was divided into a training set (≥60%) for parameter optimisation and a validation set (remainder) for testing, as detailed in Table 1.

Table 1: Characteristics and data distribution of analytes selected for PBRTQC implementation.

Analyte	Sigma value	Sample load per month	Total Data Points	Training Set (%)	Validation Set (%)
Creatinine (mg/dL)	3.47	8966	34,768	60%	40%
Urea (mg/dL)	3.10	8860	49,056	60%	40%
AST (IU/L)	4.39	6288	35,153	60%	40%
TSH (µIU/mL)	3.34	3825	13,207	69%	31%

AST: aspartate transaminase, TSH: thyroid-stimulating hormone

PBRTQC Parameter Optimization

The process to establish the PBRTQC protocol involved three systematic steps:

Truncation Limit Setting: The interquartile range (IQR) method was used to remove extreme outliers, thereby reducing skewness and stabilising the moving average calculation, The limits were calculated as: Lower Truncation Limit (LTL) = Q1 - 1.5*IQR Upper Truncation Limit (UTL) = Q3 + 1.5*IQR Moving Average and Control Limit Calculation: A Simple Moving Average (SMA) algorithm was applied: SMA = $(x_1 + x_2 + \ldots + x_n) / n$ Control limits (CLs) were set using one of two methods, chosen based on which gave a lower false alarm rate:

Method 1: The maximum and minimum of the moving average for UCL and LCL.

Method 2: The 99.95 percentile and 0.05 percentile for UCL and LCL.

Bias Simulation for Block Size Selection Introduction of bias
Different sets of bias were introduced, ranging from –50% to +50% in 10% increments, to fully characterise the bias detectioncurve across small, moderate, and large biases. In addition to this full range, bias equivalent to the CLIA-total allowable error (TEa) was specifically introduced, as the optimal block size and ANPed were determined at this clinically relevant threshold. Baseline phase: The first block of data was kept for unbiased reference. Moving average calculation: MA was calculated for the biased data set with different block sizes (10, 25, 50, 75, 100). Control limit breach: The first data point breaching UCL or LCL was identified. ANPed calculation: The average number of patient data points from bias introduction to control limit breach was recorded as ANPed. Optimal block selection: A bias detection curve (ANPed vs. Bias) was plotted. The optimal block size was selected from this curve.

Validation of Optimised Parameters
The optimised PBRTQC settings derived from the training dataset were applied to the independent validation dataset. A systematic bias equivalent to the total allowable error (TEa) from CLIA guidelines was intentionally introduced into the validation dataset to simulate analytical bias. The moving average (MA) was recalculated using the optimised block size. The point at which the MA value breached the pre-established control limits was recorded, thereby validating the error detection capability of the model.

Error Rate and Cost-Benefit Analysis
The MA alarm rate was calculated as (Number of Alarms / Total MA Data Points) × 100. The percentage improvement in error detection was calculated as (Errors detected by PBRTQC / Total Errors) × 100. A cost-benefit analysis was performed using a laboratory billing system to determine the costs of IQC runs and reagent consumption per test.

Risk-Based Cost-Benefit Analysis
The integrated implementation of PBRTQC and IQC performed a risk-based cost-benefit analysis. Benefits were mapped across four risk categories: patient safety, regulatory compliance, operational failure, and financial loss. Data sources were cross-referenced to validate each benefit.

Results

Optimised PBRTQC Parameters
The optimised PBRTQC parameters for each analyte, derived from bias simulation studies, are shown in Table 2. For creatinine, urea and TSH, the lower truncation limits were not applicable due to negative values. Block size and ANPed were derived from the bias detection curve for each analyte. When bias equal to the total allowable error (TEa) was introduced, the average number of patient data points required for error detection (ANPed) was 161 for creatinine, 207 for urea, 22 for AST, and 68 for TSH.

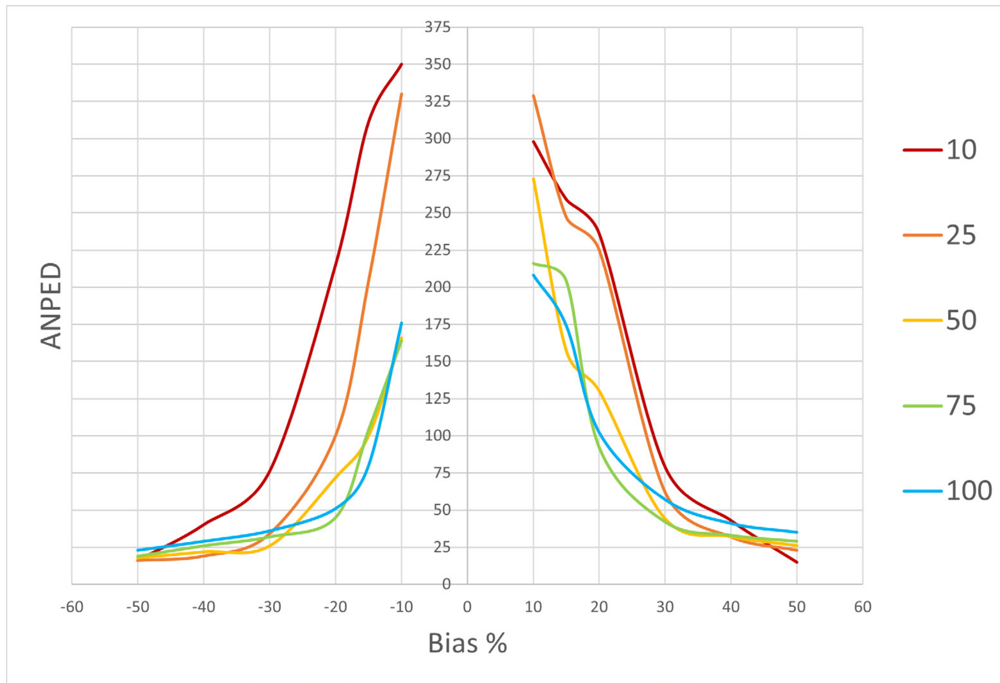
Table 2: Optimised PBRTQC parameters for creatinine, urea, AST, and TSH.

Analyte	Truncation limit	Control limit	TEa (%)	Block size	ANPed at TEa
Creatinine	UTL= 1.6 LTL= NA	UCL= 0.96 LCL= 0.59	±15%	50	161
Urea	UTL= 51 LTL= NA	UCL= 28.7 LCL= 15.9	±9%	25	207
AST	UTL= 36 LTL= 12	UCL= 26.2 LCL= 18.8	±20%	50	22
TSH	UTL= 7.8 LTL= NA	UCL= 3.7 LCL= 1.7	±20%	25	68

UTL: upper truncation limit, LTL: lower truncation limit. UCL: upper control limit, LCL: lower control limit. TEa: total error allowable, obtained from CLIA guidelines. ANPed: average number of patient data required for error detection. Block size and ANPed were obtained from bias detection curves.

Bias Detection and Validation for Creatinine
Bias detection curves were obtained by performing the bias simulation study on creatinine data. Larger biases required fewer patient data points for detection. Smaller block sizes led to falsealarms due to lower ANPeds, while larger block sizes delayed bias detection. Block sizes of 25 and 50 showed optimal performance Figure 1.

Figure 1: Bias detection curve for Creatinine.

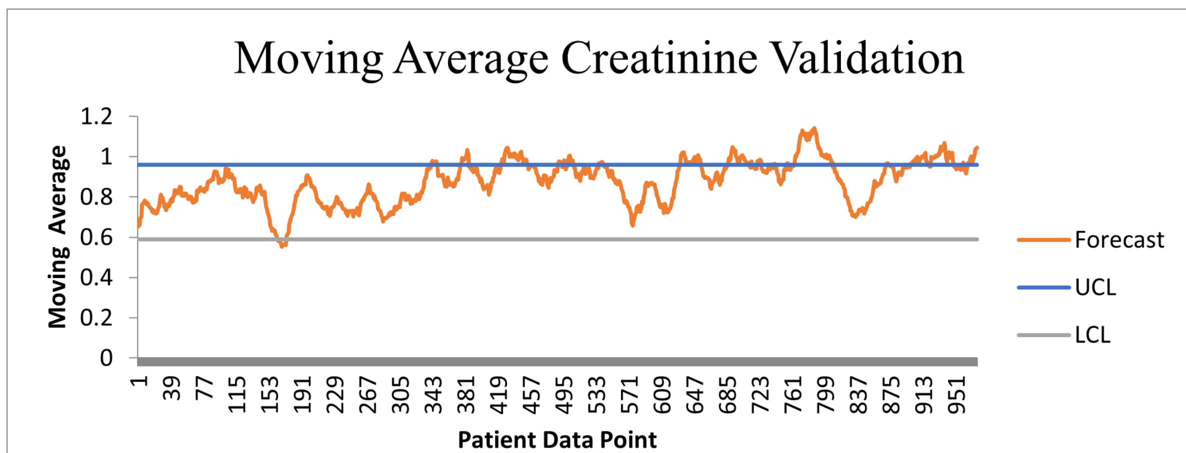


The curve shows Average Number of Patient data points required for error detection (ANPed) across varying levels of introduced systematic bias for different block sizes (10, 25, 50, 75, 100). The optimal block size of 50 was selected for its balanced performance.

Validation of the moving average algorithm for Creatinine
Average technique successfully detected the bias, as indicated by The optimised settings were validated on the validation dataset.

A the breach of control limits Figure 2. bias equivalent to CLIA TEa was introduced. The moving

Figure 2: Optimised PBRTQC parameters validation for creatinine.



**The moving average (Forecast) breaches the upper control limit (UCL) shortly after the introduction of a systematic bias equivalent to the total allowable error (TEa), confirming successful error detection. Block size = 50. **

Bias detection curves and validation charts for urea, AST, and TSH are provided in the Supplementary Material (Figures S1-S6).

Live Implementation and Alarm Analysis
The optimised moving average algorithm was applied to daily patient data forcreatinine, urea, and AST from January to March 2025, and for TSH during February and March 2025. MA alarms triggered when control limits were breached. Alarms were investigated with reanalysis of stable period patient samples.

Table 3: MA alarm investigations, root causes and corrective actions.

Analyte	Number of MA alarms generated	IQC performed (Passed/failed)	Check whether the patient was having really high / low values	Probable Cause for alarm	Corrective and Preventive Action
Creatinine (mg/dl)	UCL breach = 2 LCL breach = 3	Passed 3 times	High value = 1 time Low value = 2 times	Unclean wash nozzle leading to accumulation of debris in cuvettes Calibration not done after reagent lot change	The wash nozzle was cleaned and installed back. Monthly maintenance is to be done regularly. Calibration to be performed after every reagent lot change
Urea (mg/dl)	UCL breach = 4 LCL breach = 2	Passed 4 times	High value = 3 times Low value = 1 time	On-board reagent deterioration Unclean wash nozzle leading to accumulation of debris in cuvettes	Regular calibration should be done. The wash nozzle was cleaned and installed back
AST (IU/L)	UCL breach = 1 LCL breach = 1	Passed 2 times	High value = 1 time Low value = 1 time	Patients were having pathologically high and low values	
TSH (µIU/mL)	UCL breach = 3 LCL breach = 1	Passed 3 times	High value =1 time Low value = 1 time	Pipetting error	Pipette calibration should be done

Moving average alarm rate

Across 20143 creatinine, 22667 urea, 13776 AST and 5920 TSH moving average data points, the MA alarm percentage remainedless than 1% for all four analytes, indicating the obtained MA parameters for PBRTQC are producing less false alarm rates.

Table 4: The number of MA data points, total alarms generated, and MA alarm rates for each analyte.

Analyte	Number of MA data points	Total alarms generated	MA alarm rate (%)
Creatinine (mg/dL)	20143	5	0.0248
Urea (mg/dL)	22667	6	0.0267
AST (IU/L)	13776	2	0.0145
TSH (µIU/mL)	5920	4	0.0675

Percentage increase in error detection after PBRTQC and IQC integration

IQC and PBRTQC data were collected over three months (January 2025 to March 2025) for creatinine, urea, and AST, and for TSH during February 2025 and March 2025. The percentage improvement in error detection was calculated for the integration of IQC and PBRTQC.

Table 5: shows PBRTQC integration improved error detection for Creatinine (28.57%), Urea (25%), and TSH (20%), Consistent with its role as a comparator, AST the only analyte with a sigma value above 4 showed no improvement in error detection with PBRTQC integration.

Analyte	No of errors detected	No of errors detected	No of errors detected	% increase in error detection after IQC-PBRTQC integration
Analyte	IQC	PBRTQC	Total	% increase in error detection after IQC-PBRTQC integration
Creatinine	5	2	7	28.57%
Urea	6	2	8	25%
AST	3	0	3	0%
TSH	4	1	5	20%

Both PBRTQC alarms for creatinine and urea, and one alarm for TSH, were exclusive error detections occurring between scheduled IQC runs, confirmed by reanalysis of stable-period samples. These represent errors that would have remained undetected until the next IQC run. AST: aspartate transaminase, TSH: thyroid-stimulating hormone

Cost-Benefit Analysis

The scheduled IQC frequency remained unchanged. The reported cost savings reflect the unscheduled IQC runs that would havebeen required to detect errors occurring between scheduled QC intervals in the absence of PBRTQC. The integration model yielded a net saving of INR 426.92 from avoided IQC runs and reanalysis costs.

Table 6: Potential Cost-benefit analysis of PBRTQC implementation**.**

Analysis Type	Creatinine	Urea	TSH	Total
Errors detected by PBRTQC	2	2	1	5
Cost per IQC run (INR)	15.90	26.38	180.00	
IQC savings (INR)	31.80	52.76	180.00	264.56
Cost per reanalysis (INR)	9.10	22.08	100.00	
Reanalysis savings (INR)	18.20	44.16	100.00	162.36
Total savings (INR)	50.00	96.92	280.00	426.92

Savings represent the unscheduled IQC runs that would have been needed to detect errors occurring between scheduled QC intervals. IQC savings = Number of errors detected by PBRTQC × Cost per IQC run Reanalysis savings = Number of errors detected by PBRTQC × Cost per reanalysis Total savings = IQC savings + Reanalysis savings

Risk-based Cost-benefit advantage

Table 7 demonstrates the synergistic benefits of integrating PBRTQC with conventional IQC across four risk domains. Key findings include a 20–28.6% improvement in error detection (enhancing patient safety), compliance with CLIA/ISO 15189 through continuous monitoring, proactive identification of operational failures (e.g., reagent deterioration), and net cost savings of 426.92 (Table 6). Data sources are cross-referenced to validate each benefit.

Table 7: Risk-based Cost-benefit advantage.

Risk category	Key benefit of PBRTQC	Supporting data from the study
Patient safety	20 to 28.6% faster error detection rate vs IQC alone	Table 5 (% increase in error detection)
Regulatory compliance	Continuous monitoring meets CLIA/ ISO 15189	Table 2 (ANPed calculation at bias equal to TEa as per CLIA guidelines)
Operational failure	Early detection of errors (reagent/calibration issues)	Table 3 (alarm and root cause analysis)
Financial loss	Saves IQC material and avoids reanalysis	Table 6 (net cost benefit)

CLIA: clinical laboratory improvement amendments, IQC: internal quality control, ISO:International organization for standardization

Discussion

The implementation of Patient-Based Real-Time Quality Control (PBRTQC) in our clinical biochemistry laboratory using Microsoft Excel was guided by three key prerequisites: (1) prioritizing analytes with lower sigma values, as these parameters are inherently more error-prone and thus require heightened monitoring; (2) accommodating the laboratory's high monthly sample volume; and (3) leveraging robust historical datasets to ensure methodologically sound parameter optimization.

Parameter Selection and Data Robustness

We prioritised analytes with lower sigma values because they are more susceptible to analytical errors, making their continuous monitoring crucial for quality assurance. The use of historical data allowed us to simulate realistic lab conditions, ensuring that the PBRTQC settings derived from this training dataset were robust and practical for actual use.

An essential step in the PBRTQC setup was outlier handling. We adopted an interquartile range (IQR)-based truncation approach to eliminate outliers, thereby reducing skewness of the dataset. This helped reduce the influence of extreme values and ultimately decreased several false alarms. Although winsorization could be a more effective technique, leading to earlier error detection, as highlighted by Badrick et al [9]. Our IQR method still produced satisfactory results.

Control Limit Optimisation

With the help of real historical data, tight control limits were established, which proved to be critical for early error detection. Tightening the limits and reducing false alarms was found to be done using a percentile of daily extremes (specifically the 0.05 and 99.95 percentiles), which was more effective, and this particular approach was very consistent with the findings from the work of Van Rossum [2,10]. Optimal block size was determined with the bias simulation study using the bias detection curves and the average number of patient data points required for error detection (ANPed). Early error detection was observed with the smaller block sizes, which in turn led to an increase in the rate of false alarms, pointing towards the need for balance.

Alarm Management Strategy

Establishing a particular alarm management strategy was done after selecting the optimal Moving Average (MA) parameters. According to our findings, individual pathological findings oranalytical errors have shown a significant proportion of alarms. This validated the practice of reviewing patient results before the alarm [9].

In our laboratory, upon each PBRTQC alarm, stable-period patient samples were immediately reanalysed. If bias was confirmed, IQC was performed immediately for double-confirmation and root cause analysis. If reanalysis showed no bias, IQC was performed at its scheduled interval, with subsequent reanalysis confirming the values reflected true pathological highs or lows. This two-step verification protocol aligns with recommendations as suggested by Badrick et al. [9,11].

In our study, alarm fatigue was not observed with the frequency of alarms; instead, the number of alarms was seen to remain manageable.

Validation, practical impact, and cost benefit

Since AST demonstrated relatively better sigma performance compared to the other three analytes, it showed fewer MA alarms. Consistently, the absence of benefit for AST supports the specificity of our observations and suggests that PBRTQC integration yields maximum value for assays with sigma values below the commonly accepted threshold for acceptable performance.

For validation, we introduced bias equivalent to the total allowable error (TEa), as recommended by the IFCC working group on PBRTQC [12]. The optimized settings successfully detected this bias, confirming their reliability. Our previous sigma matrix evaluations, conducted using CLIA guidelines, were consistent with the TEa values used in this study [13].

The clinical utility of ANPed as a performance metric must be interpreted in the context of the laboratory's sample volume. A higher ANPed does not necessarily equate to delayed error detection if daily sample throughput is proportionally high, as demonstrated with urea in this study. This reinforces the importance of laboratory-specific parameter optimisation rather than the adoption of generic block sizes.

The integration of PBRTQC with IQC improved error detection for low-sigma analytes while also demonstrating cost benefits. PBRTQC complements IQC by providing an additional layer of continuous error detection during intervals between scheduled QC runs. The cost savings reported reflect the unscheduled IQC runs avoided - those that would have been required to detect

theseerrors without PBRTQC - demonstrating a cost-benefit advantage without altering the existing QC schedule. By enabling earlier detection of analytical errors, this integrated approach reduces patient risk by preventing the release of erroneous results that could lead to misdiagnosis.
Furthermore, our study bridges regulatory and quality management frameworks: validation against CLIA's TEa meets analytical requirements, while the risk-based approach aligns with ISO 15189 goals [14,15].

Conclusion

The study demonstrated that Patient-Based Real-Time Quality Control (PBRTQC) can be effectively implemented using simple tools like Excel to monitor laboratory analytes, especially with low sigma performance. By optimising block sizes, control limits, and truncation methods, PBRTQC successfully detected analytical errors early with minimal false alarms. Integration with traditional IQC further strengthened error detection, improving overall laboratory quality and patient safety cost-effectively and practically. This integrated approach provides a strong risk-based cost-benefit advantage, enhancing compliance with international standards while proactively mitigating patient risk through continuous quality monitoring.

Limitations

In this study, a limited number of analytes were evaluated for PBRTQC procedures.
Simple Moving Average (SMA) was used; employing an Exponentially Weighted Moving Average (EWMA) could have been more effective.
Paediatric creatinine patient data were not excluded, which may have influenced the results due to physiologically lower creatinine values in children.
A limitation of this study is that the error events detected by IQC and PBRTQC were not mutually exclusive, precluding formal statistical comparison. The findings are therefore presented as descriptive improvements.

Acknowledgements

The authors would like to extend their sincere gratitude to **Naveen Ahmad Nadaf, Ather Paul, and Dr Sandhya R** for their valuable suggestions and helpful ideas during the initial stages of this project.

Author Statements

Credit Author Statement

We'd like to clarify how each author contributed to this work:
Mudasir Bashir Dandroo: Was responsible for the core ideas, performing the analysis, working with the data, writing the first draft, and creating the figures.
Dr Devanatha Desikan V: Assisted in data collection, contributed to cost benefit analysis, provided valuable suggestions during study design and offered overall study oversight.
Dr Ramesh R: Provided the initial concept and designed the study, provided guidance, oversight throughout the project, resources, and thoroughly reviewed and edited the manuscript
As the corresponding author, I confirm that this accurately reflects our contributions.

Declaration of Competing Interests

The authors declare that there is no conflict of interest.

Ethical Approval

Our study used existing, anonymised lab data. The **Ethics Committee at JIPMER, Puducherry,** reviewed and approved this project (Ref: JIP/IEC-OS/2023/053). Because the data was historical and anonymous, the committee waived the requirement for individual patient consent.

Funding Statement

This work was carried out as part of our academic research. No external funding from grants or companies was received for this specific study.

Data Availability Statement

The anonymised patient data that support the findings of this study are available from the corresponding author, upon reasonable request and with permission from the Institutional Ethics Committee of JIPMER.

References

1. Gruber L, Hausch A, Mueller T. Internal Quality Controls in the Medical Laboratory: A Narrative Review of the Basic Principles of an Appropriate Quality Control Plan. Diagnostics (Basel). 2024 Oct 5;14(19):2223. doi:10.3390/diagnostics14192223.
2. van Rossum HH. Moving average quality control: principles, practical application and future perspectives. Clin Chem Lab Med. 2019;57(5):773-782. doi:10.1515/cclm-2018-0794.
3. van Rossum HH, van den Broek D. Design and implementation of quality control plans that integrate moving average and internal quality control: incorporating the best of both worlds. Clin Chem Lab Med. 2019;57(9):1329-1338. doi:10.1515/cclm-2018-1306.
4. Hinge M, Lund ED, Brandslund I, Madsen JS. Patient pools and the use of "patient means" are valuable tools in quality control,illustrated by a bone-specific alkaline phosphatase assay. Clin Chem Lab Med. 2016;54(3):403-406. doi:10.1515/cclm-2015-0696.
5. Hoffmann RG, Waid ME. The "Average of Normals" Method of Quality Control. Am J Clin Pathol. 1965;43:134-141. doi:10.1093/ajcp/43.2.134.
6. van Rossum HH, Kemperman H. Moving average for continuous quality control: time to move to implementation in daily practice? Clin Chem. 2017;63(6):1041-1043. doi:10.1373/clinchem.2016.269027.
7. Schoenmakers CH, Naus AJ, Vermeer HJ, van Delft JP, van Oosterhout M. Practical application of sigma metrics QC procedures in clinical chemistry. Clin Chem Lab Med. 2011;49(11):1837-1843. doi:10.1515/CCLM.2011.681.
8. Westgard JO, Bayat H, Westgard SA. Planning risk-based SQC schedules for bracketed operation of continuous

production analysers. Clin Chem. 2018;64(2):289-296. doi:10.1373/clinchem.2017.278135.

9. Badrick T, Bietenbeck A, Katayev A, van Rossum HH, Loh TP, Cervinski MA, et al. Implementation of patient-based real-time quality control. Crit Rev Clin Lab Sci. 2020;57(8):532-547. doi:10.1080/10408363.2020.1765731.
10. Bietenbeck A, Cervinski MA, Katayev A, Loh TP, van Rossum HH, Badrick T. Understanding Patient-Based Real-Time Quality Control Using Simulation Modelling. Clin Chem. 2020;66(8):1072-1083. doi:10.1093/clinchem/hvaa094.
11. Duan X, Zhang M, Liu Y, Zheng W, Lim CY, Kim S, Loh TP, Guo W, Zhou R, Badrick T, et al.; Patient-Based Real-Time Quality Control Working Group of the Asia Pacific Federation of Clinical Biochemistry and Laboratory Medicine. Next-generation patient-based real-time quality control models. Ann Lab Med. 2024;44(5):385–391. Epub 2024 Jun 5. doi: 10.3343/alm.2024.0053. PubMed PMID: 38835211; PubMed Central PMCID: PMC11169771.
12. Loh TP, Cervinski MA, Katayev A, Bietenbeck A, van Rossum H, Badrick T, et al. Recommendations for laboratory informatics specifications needed for the application of patient-based real-time quality control. Clin Chim Acta. 2019;495:625-629. doi:10.1016/j.cca.2019.06.009.
13. Rohit S, Ramesh R. Effect of Internal Quality Planning using Sigma Metrics in Lean Management of a Clinical Chemistry Laboratory: An Analytical Study. J Clin Diagn Res. 2024;18(4):BC12-BC17. doi:10.7860/JCDR/2024/67309/19260.
14. Palmer J, Galior K. Defining allowable total error limits in the clinical laboratory. Adv Clin Chem. 2024;118:205-223. doi:10.1016/bs.acc.2023.11.006.
15. Badrick T, Loh TP. Integrating patient-based quality control and patient-based quality assurance. Clin Biochem. 2024;124:110708. doi:10.1016/j.clinbiochem.2024.110708.

Supplementary files

PBRTQC performance characteristics – Urea

For urea same methodology which was employed for creatinine was used for urea, to obtain optimum truncation limits, control limits and block size and ANPed were calculated by performing a bias simulation study and plotting bias detection curves.

Table 1: Shows the PBRTQC performance characteristics for urea, including truncation limits, control limits, TEa, block size, and ANPed.

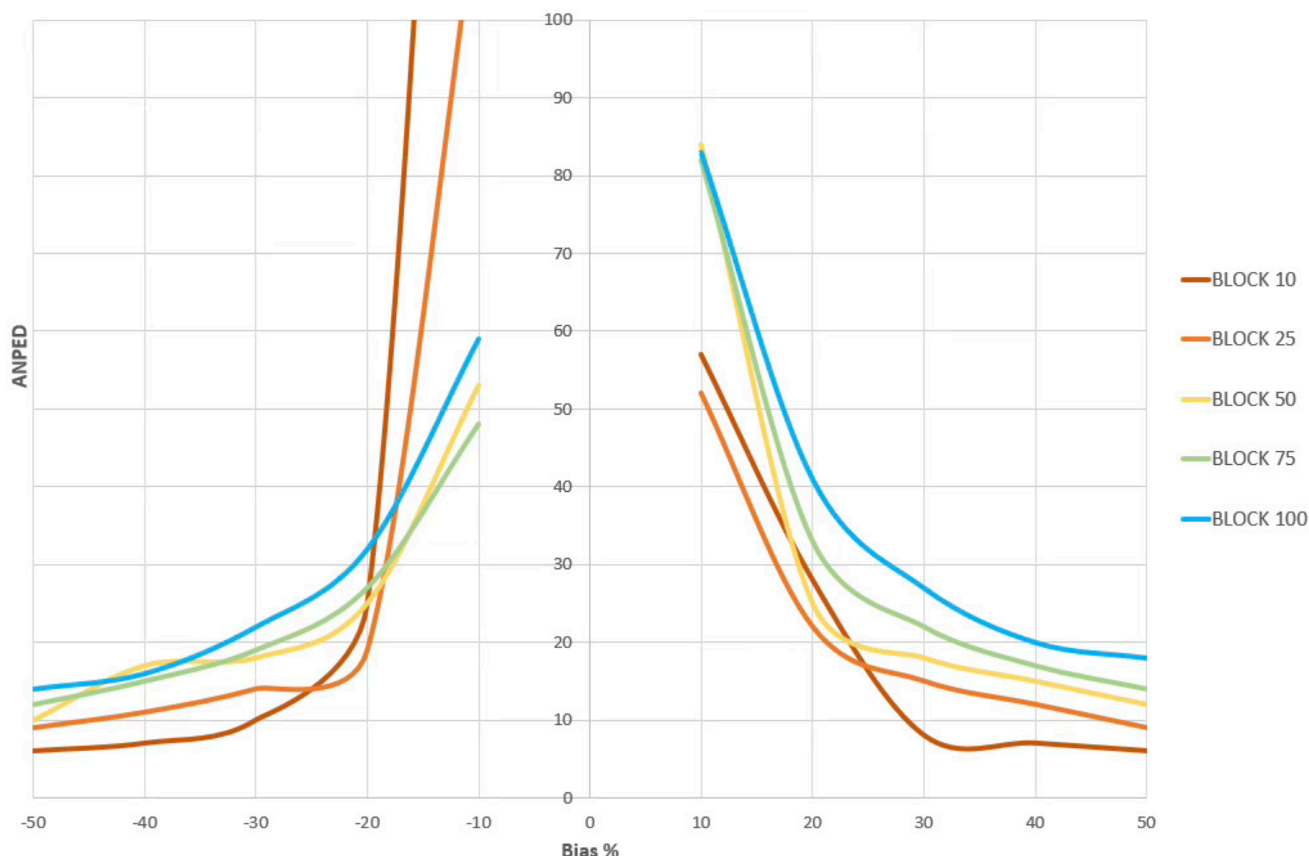
Analyte	Truncation limit	Calculation algorithm	Control limit	Tea (%)	Block size*	ANPed* At TEa
Urea	UTL= 51 LTL= NA	SMA	UCL= 28.7 LCL= 15.9	±9%	25	207

UTL: upper truncation limit, LTL: lower truncation limit. UCL: upper control limit, LCL (lower control limit). TEa: total error allowable, it was obtained from the CLIA guidelines. ANPed: average number of patient data required for error detection
**Block size and ANPed were obtained from bias detection curves*

Bias detection curves were obtained by performing the bias simulation study on urea

Larger biases required fewer patient data points for detection. Smaller block sizes (e.g., 10) led to false alarms due to lower ANPeds, while larger block sizes (75-100) delayed bias detection due to higher ANPeds requirements. Block size of 25 with ANPed 207 at error equal to TEa showed optimal performance.

Figure 1A: Optimizing PBRTQC Using Bias Detection Curves.

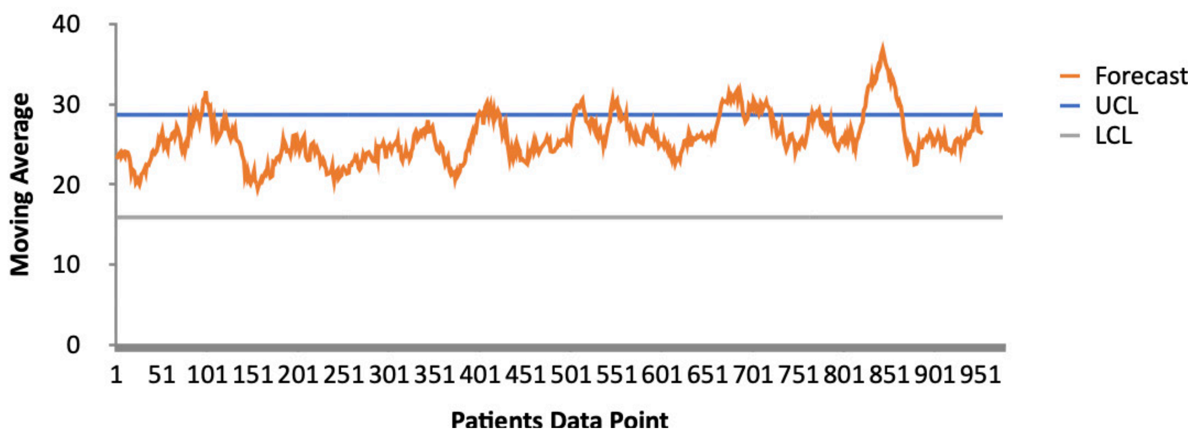


This figure compares urea PBRTQC protocols using the mean of the most recent 10, 25, 50, 75, and 100 patient results. It shows the average number of patient samples required to detect an error (ANPed) across varying systematic bias levels (-50% to +50%), enabling direct comparison to identify the most effective error detection protocol.

Validation of derived optimized parameter settings on the validation data set

Optimized PBRTQC settings were derived from a urea training dataset and applied to a 1,000-point validation dataset with simulated analytical bias (TEa). The moving average (MA) technique detected the bias via a control limit breach.

Figure 1B: Shows the moving average (red line) breaches upper control limit, shortly after the introduction of systematic bias equivalent to total allowable error, indicating the successful detection of bias, Block used to calculate moving average was kept equal to 25.



PBRTQC performance characteristics - AST

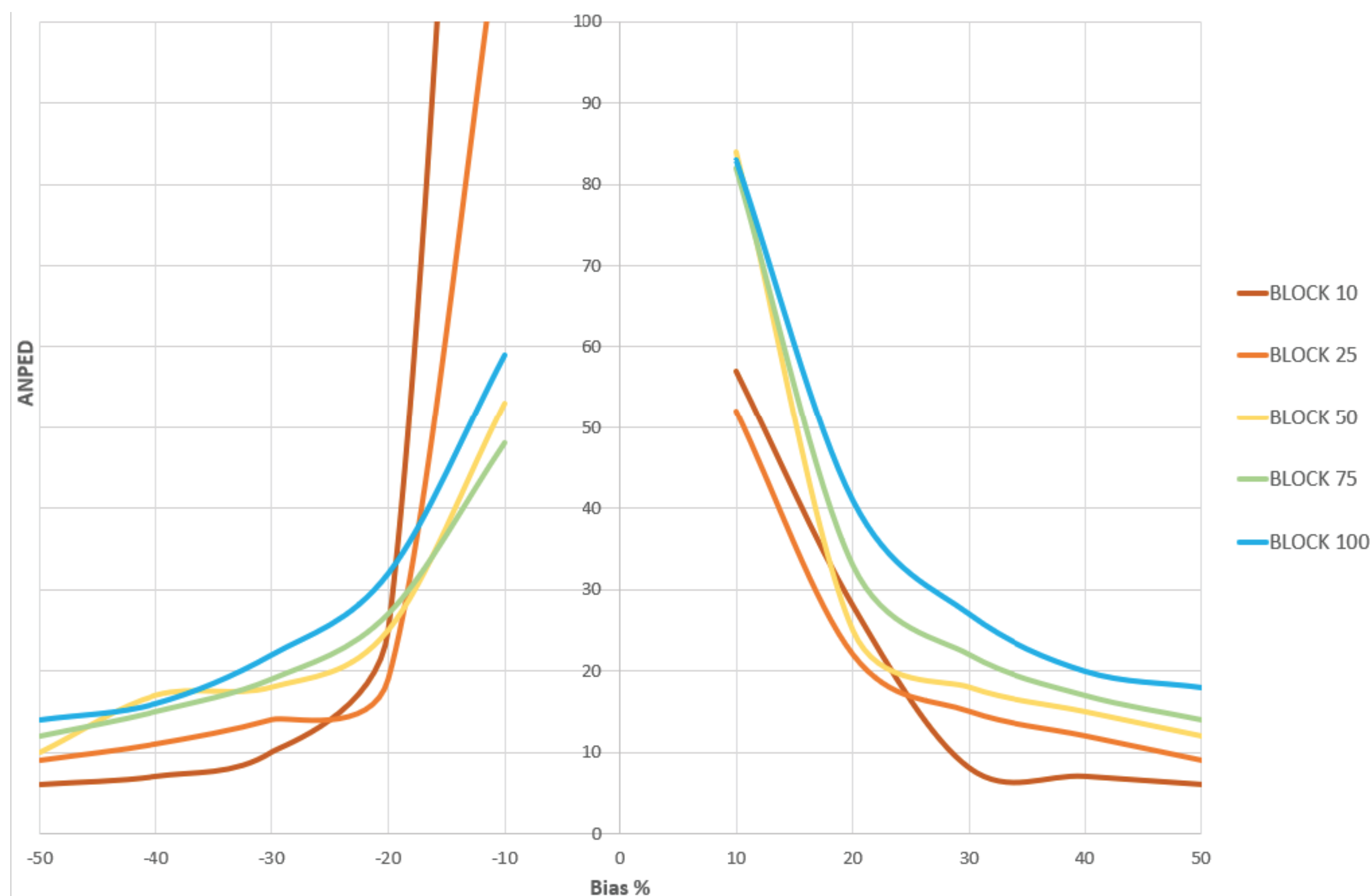
Similar methodology was used to AST, to obtain optimum truncation limits, control limits and block size and ANPed were calculated by performing a bias simulation study and plotting bias detection curves.

Table 2: shows the PBRTQC performance characteristics for AST, including truncation limits, control limits, TEa, block size, and ANPed.

Analyte	Truncation limit	Calculation algorithm	Control limit	TEa (%)	Block size*	ANPed* At TEa
AST	UTL= 36 LTL= 12	SMA	UCL= 26.2 LCL= 18.8	±20%	50	22

UTL: upper truncation limit, LTL: lower truncation limit. SMA: simple moving average UCL: upper control limit, LCL (lower control limit). TEa: total error allowable, it was obtained from the CLIA guidelines. ANPed: average number of patient data required for error detection
**Block size and ANPed were obtained from bias detection curves*

Figure 2A: Optimizing PBRTQC Using Bias Detection Curves.



This figure compares AST PBRTQC protocols using the mean of the most recent 10, 25, 50, 75, and 100 patient results. It shows the average number of patient samples required to detect an error (ANPed) across varying systematic bias levels (-50% to +50%), enabling direct comparison to identify the most effective error detection protocol

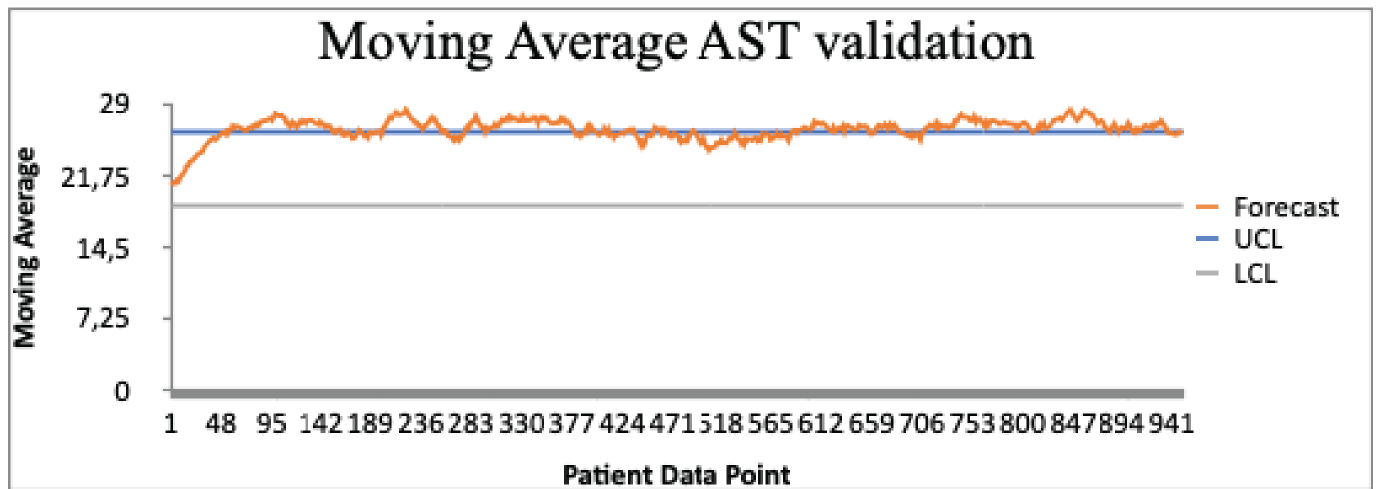
Bias detection curves were obtained by performing the bias simulation study on AST

larger biases required fewer patient data points for detection. Smaller block sizes (e.g., 10) led to false alarms due to lower ANPeds, while larger block sizes (75-100) delayed bias detection due to higher ANPeds requirements. Block sizes of 50 with ANPed 22 at error equal to TEa showed optimal performance.

Validation of derived optimized parameter settings on the validation data set

Optimized PBRTQC settings were derived from a urea training dataset and applied to a 1,000-point validation dataset with simulated analytical bias (TEa). The moving average (MA) technique detected the bias via a control limit breach.

Figure 2B: Shows the moving average (red line) breaches upper control limit, shortly after the introduction of systematic bias equivalent to total allowable error, indicating the successful detection of bias Block used to calculate moving average was kept equal to 50.



PBRTQC performance characteristics
Similar methodology as described above was used for TSH, to obtain optimum truncation limits, control limits and block size and ANPed were calculated by performing a bias simulation study and plotting bias detection curves.

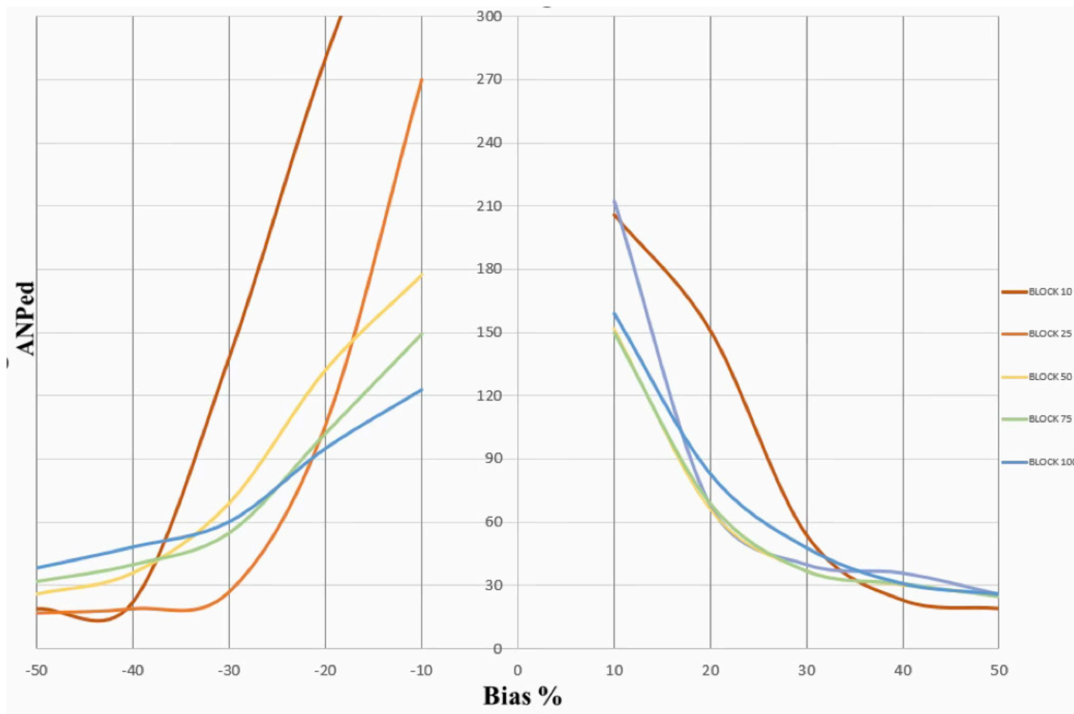
Table 3: PBRTQC performance characteristics for TSH, including truncation limits, control limits, TEa, block size, and ANPed.

Analyte	Truncation limit	Calculation algorithm	Control limit	TEa (%)	Block size*	ANPed* At TEa
TSH	UTL= 7.8 LTL= NA	SMA	UCL= 3.7 LCL= 1.7	±20%	25	68

UTL: upper truncation limit, LTL: lower truncation limit. SMA: simple moving average UCL: upper control limit, LCL (lower control limit). TEa: total error allowable, it was obtained from the CLIA guidelines. ANPed: average number of patient data required for error detection
**Block size and ANPed were obtained from bias detection curves*

Bias detection curves were obtained by performing the bias simulation study on TSH
larger bias percentages required fewer patient data points for detection. Smaller block sizes (e.g., 10) led to false alarms due to lower ANPeds, while larger block sizes (75-100) delayed bias detection due to higher ANPeds requirements. Block sizes of 25 with ANPed 68 at error equal to TEa performed best.

Figure 3A: Represents Optimization of PBRTQC Using Bias Detection Curves: This figure compares urea PBRTQC protocols using the mean of the most recent 10, 25, 50, 75, and 100 patient results.

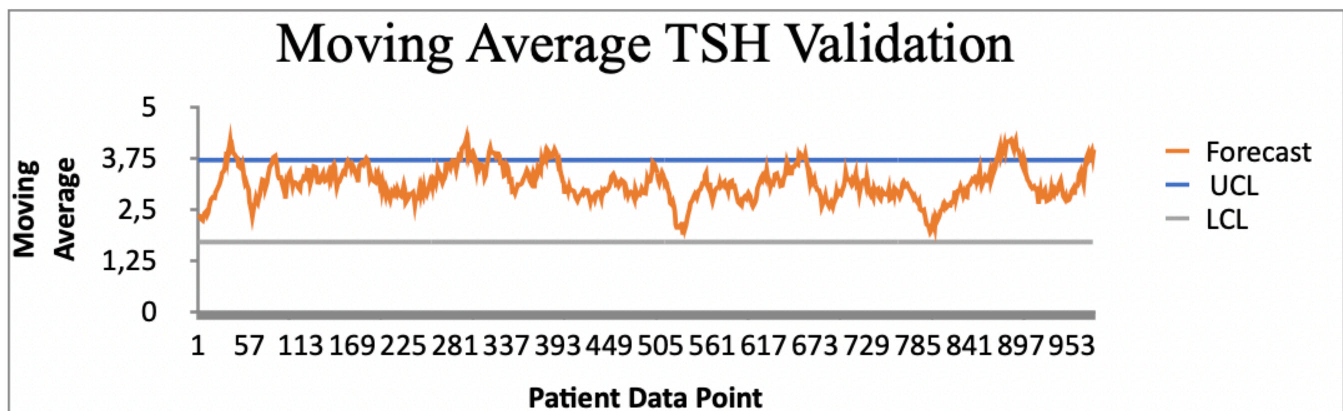


It shows the average number of patient samples required to detect an error (ANPed) across varying systematic bias levels (-50% to +50%), enabling direct comparison to identify the most effective error detection protocol.

Validation of derived optimized parameter settings on validation data set

Optimized PBRTQC settings were derived from a urea training dataset and applied to a 1,000-point validation dataset with simulated analytical bias (TEa). The moving average (MA) technique detected the bias via a control limit breach.

Figure 3B: Shows the moving average (red line) breaches upper control limit, shortly after the introduction of systematic bias equivalent to total allowable error, indicating the successful detection of bias Block used to calculate moving average was kept equal to 25.



Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Research Article

The College of American Pathologists (CAP) Laboratory Accreditation Program (LAP): Accreditation Process and its Impact on Laboratory Practice and Patient Safety

Abdelaziz Sanad Abdelaziz[1,2], Samir Y. Marzouk[2,3], and Mai S. Mabrouk[4*]

[1]*Advanced Cell Laboratories, Riyadh, KSA.*
[2]*Productivity and Quality Institute (PQI), Arab Academy for Science, Technology, and Maritime Transport, Cairo, Egypt.*
[3]*Basic and Applied Science Department, College of Engineering and Technology, Arab Academy for Science, Technology, and Maritime Transport, Cairo, Egypt*
[4]*Center for Information (CIS), School for Information Technology and Computer, Nile University, 26th Of July Corridor, Sheikh Zayed City, Giza, Egypt*

Article Info

Corresponding Author:
Mai S. Mabrouk
Center for Information (CIS), School for Information Technology and Computer, Nile University, 26th Of July Corridor, Sheikh Zayed City, Giza, 12588, Egypt
E-mail: masaid@nu.edu.eg
Phone: +201001662403

Keywords
Accreditation impact, CAP LAP, Laboratory Accreditation Program (LAP), laboratory quality, patient safety, and quality management

Abstract

This study aimed to assess the laboratory professional perceptions on how the College of American Pathologists (CAP) Laboratory Accreditation Program (LAP) is associated with the laboratory processes and patient safety across multiple countries in the Middle East and South Asia. The objective was to determine how CAP LAP influences laboratory quality and safety culture, and participants reported that error-related events appeared to decrease following the CAP LAP. A quantitative, cross-sectional design was used, employing a structured electronic questionnaire distributed to CAP-accredited laboratories in Saudi Arabia, other Gulf countries, Egypt, and Pakistan. A total of 300 valid responses were analyzed using descriptive statistics and marginal homogeneity testing to compare the respondents' perceived condition pre- and post-accreditation performance in areas such as staff competency, quality management, turnaround time, and safety practices. Respondents assessed conditions before and after accreditation in pre-analytical, analytical, and post-analytical areas, including quality management systems, staff competence, turnaround time, equipment upkeep, communication, and safety culture. We examined paired ordinal responses using marginal homogeneity tests. The CAP LAP was therefore viewed as highly valuable for improving both laboratory performance and patient safety. This study provides empirical evidence from multi-country CAP LAP accredited laboratories, filling a regional gap in accreditation research. It demonstrates that the CAP LAP functions as a strong governance mechanism enhancing quality and safety within diverse healthcare systems. It is advisable to conduct future longitudinal and comparative studies among CAP, ISO 15189:2022, and national accreditation systems to evaluate the durability of enhancements and their impact on laboratory personnel behavior. These results indicate that laboratory professionals view CAP LAP as linked to significant enhancements in laboratory operations and safety protocols. However, the results are based on people's own reports of their perceptions and don't show cause-and-effect relationships or measure objective safety outcomes. Future research incorporating objective performance metrics and longitudinal methodologies is essential to assess the causal influence of accreditation on patientsafety.

Introduction

Clinical laboratories are the cornerstone of modern healthcare, supporting approximately 70% of medical decisions related to diagnosis, patient monitoring, and prognosis [1]. The laboratory testing process includes three key phases: pre-analytical, analytical, and postanalytical. Each phase involves specific procedures, with unique challenges and risks for potential errors [2]. Accreditation programs appear to improve the structure and process of care, with a substantial body of evidence showing that they improve clinical outcomes. Accreditation programs in medical laboratories such as ISO 15189 and other international and national laboratory accreditation programs emphasizes risk-based approaches and patient safety [3-6]. In 1961, CAP started the Laboratory Accreditation Program (LAP) as one of the most internationally recognized laboratory accreditation programs to provide labs a strict, peer-based inspection accreditation program addresses all the laboratory areas, including directorship, quality control, test methodologies, reagents, control media, equipment, specimen handling, procedure manuals, test reporting, proficiency testing and monitoring, as well as personnel safety and overall management practices that distinguish a quality laboratory [7-10]. Although the benefits of CAP LAP are well-documented in terms of overall laboratory performance, limited research has explored the specific perceived impact of accreditation on laboratory processes (preanalytical, analytical, and postanalytical). There is a regional knowledge gap in understanding the perceived influence of CAP LAP standards in the Middle east region [9]. Therefore, the primary objective of this study is to assess the perceived impact of the CAP LAP on laboratory procedures, professional competencies, and patient safety outcomes across CAP LAP accredited laboratories in Saudi Arabia, Gulf countries, Egypt, and Pakistan.

Materials and Methods

Study design

The study used a quantitative cross-sectional survey design. A self-administered questionnaire was distributed electronically to laboratory staff working in CAP LAP accredited laboratories. Respondents were asked to compare the laboratory practices prior to accreditation and after achieving CAP LAP accreditation [19].

Conceptual Framework

The study hypothesized that CAP LAP as independent factor positively influences laboratory process and improve laboratory process and patient safety as dependent factor [19]. The proposed conceptual framework follows the following pathway: CAP LAP → Quality management and governance mechanisms (mediator) → Lab Processes → Patient Safety outcome.

Setting and Participant

The participants include laboratory directors, quality managers, section heads, supervisors, laboratory technologists and technicians, and other senior lab staff across Saudi Arabia (the main focus), other Gulf countries, Egypt, and Pakistan. The inclusion criteria were laboratory professionals (directors, quality managers, supervisors, and technologists) working in CAP-accredited laboratories; individuals with at least one year of laboratory experience. The exclusion criteria encompass staff from laboratories lacking CAP LAP, administrative personnel not directly engaged in laboratory operations or patient safety, and incomplete or duplicate survey responses. A Sample of 300 respondents completed the survey.

Survey Instrument

A structured electronic survey with 29 questions was developed based on the following themes

- Demographics and laboratory profile (role, lab type, years accredited, staff size, hospital-based or standalone laboratory).
- Perceptions prior to CAP (inspection, QMS, staff competency, error reporting, safety, LIS, communication).
- Perceptions subsequent to CAP (identical indicators for direct comparison).
- Impact on patient safety (effectiveness of error reporting, reduction of near-misses, culture, corrective actions, communication).
- Evaluation of the accreditation process (preparedness, training, clarity of CAP standards, inspector support, challenges).
- Overall outcomes value of CAP LAP, enhancements in quality and safety.

Questions include Multiple-choice questions (role, lab type, years accredited), Likert scale (1–5, strongly disagree → strongly agree), linear rating scales (1 = very poor → 5 = excellent), before vs. after comparisons, and checkboxes for multiple areas of improvement/challenges.

Survey Instrument Development and Validity

The survey was created by looking at the literature on laboratory accreditation, quality management systems, and patient safety in medical labs, with a focus on the requirements of the CAP Laboratory Accreditation Program. A preliminary pool of items was created to address important areas, such as quality management, laboratory process performance during the pre-analytical, analytical, and post-analytical phases, safety culture, and patient safety outcomes. A pilot test with (15–20) laboratory professionals from (CAP LAP) labs was conducted to see how clear, understandable and relevance it was to fill out. Based on feedback minor revisions were made to question wording and response scales to improve clarity, consistency, and ease of completion.

Internal consistency Reliability was assessed using (Cronbach's alpha) and the values were: Pre-(CAP LAP): $\alpha = 0.93$ (Excellent) Post-CAP LAP: $\alpha = 0.92$ (Excellent), Improvement Scale: $\alpha = 0.88$ (Good) and Safety Culture Scale: $\alpha = 0.87$ (Good), the results of reliability tests show strong psychometric reliability of the instrument. All domains exhibited satisfactory to exceptional internal reliability ($\alpha = 0.78$–0.91), Internal consistency was assessed using Cronbach's alpha ($\alpha \geq 0.70$).

Data Collection Procedures

The survey was distributed electronically between August and October 2025. Participation was voluntary and anonymous; the

fully completed responses were included in the database.

Ethical Considerations

The study involved anonymous survey responses from laboratory professionals and didn't include patient data or clinical intervention. The study involved minimal-risk survey data from professionals. Implied informed consent was obtained, and confidentiality was maintained.

Statistical Analysis

The data was analyzed using the Statistical Package for Social Science (SPSS 27) and Python software. Descriptive statistics: Frequency and percentage were used for categorization of the variables. Marginal homogeneity test assesses the statistical significance of the difference of a variable with multiple categories measured. The P-value represented the level of significance P>0.05: Non-significant (NS). P < 0.05: Significant (S). N/A responses were treated as missing and excluded pairwise from the analysis.

Results

A total of 300 professionals working in CAP-accredited medical laboratories in Saudi Arabia, other Gulf countries, Egypt, and Pakistan participated in the survey. The results are presented in the following sections

Characteristics of Participating Laboratories

The general data assessment shows that among participants, 37% were laboratory directors, followed by 22% technologists/technicians, 20% section heads/supervisors, 17% quality managers, and only 4% in other roles. Most laboratories were independent/private (72%), while 28% were hospital-based. Regarding CAP LAP, more than half (54%) had been accredited for over 6 years, while 24% reported 1–3 years, 16% reported 4–6 years, and only 6% had less than 1 year. In terms of staff size, 52% of laboratories had between 20 and 50 employees, 38% had more than 50, while only 10% operated with fewer than 20 staff members The demographic characteristics of the participating laboratories are summarized in Table 1.

Table 1: Demographic Characteristics of Participating Laboratories.

Variables	Category	N (%)
Participants role in the laboratory	Laboratory Director	111 (37%)
Participants role in the laboratory	Laboratory Technologist/Technician	66 (22%)
Participants role in the laboratory	Quality Manager	51 (17%)
Participants role in the laboratory	Section Head/Supervisor	60 (20%)
Participants role in the laboratory	Other	12 (4%)
Type of laboratory	Hospital-based	84 (28%)
Type of laboratory	Independent/private	216 (72%)
Number of years the lab has been CAP-accredited	Less than 1 year	18 (6%)
Number of years the lab has been CAP-accredited	1–3 years	72 (24%)
Number of years the lab has been CAP-accredited	4–6 years	48 (16%)
Number of years the lab has been CAP-accredited	More than 6 years	162 (54%)
Approximate number of staff in the lab	Less than 20	30 (10%)
Approximate number of staff in the lab	20–50	156 (52%)
Approximate number of staff in the lab	More than 50	114 (38%)

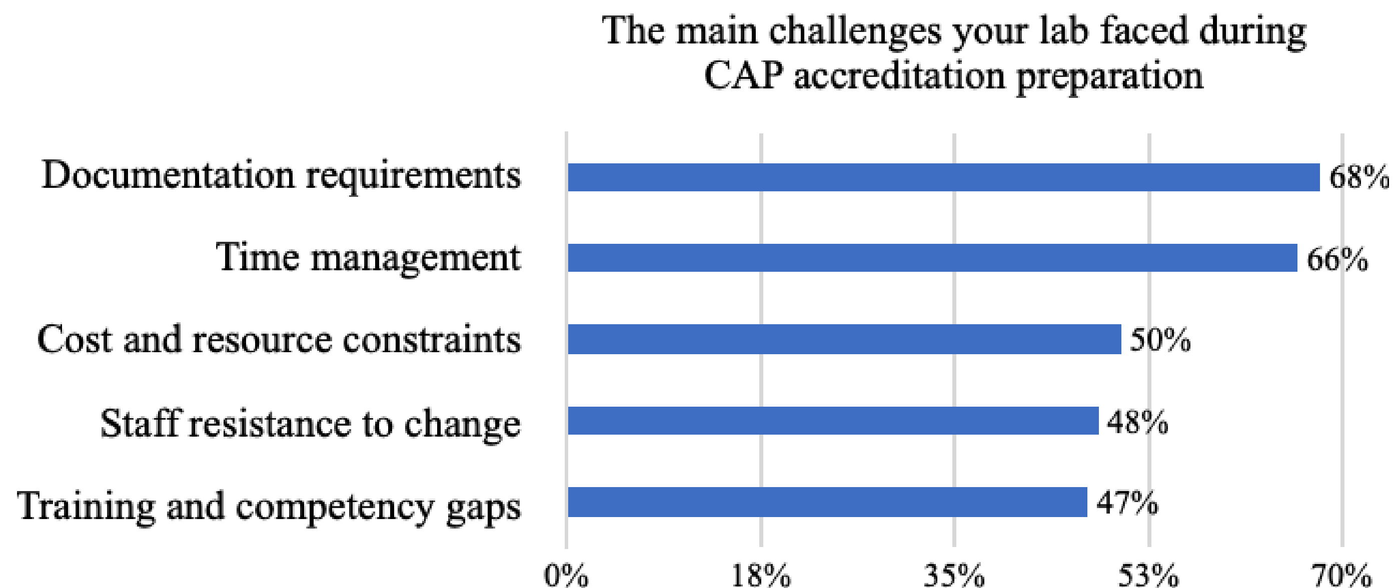
Main Challenges During CAP LAP Accreditation Preparation

The main challenges during CAP LAP preparation were dominated by documentation requirements (68%) and time management (66%), followed closely by cost/resource constraints (50%) and staff resistance to change (48%). Training and competency gaps were also a significant barrier (47%). These findings indicate that the administrative workload and time pressure associated with CAP LAP accreditation implementation. The main challenges during CAP LAP accreditation preparation are summarized in Table 2 and Figure 1.

Table 2: The main challenges the lab faced during CAP LAP preparation.

Variable	N (%)
Staff resistance to change	144 (48%)
Documentation requirements	204 (68%)
Time management	198 (66%)
Cost and resource constraints	150 (50%)
Training and competency gaps	141 (47%)

Figure 1: The main challenges faced by the laboratory during CAP LAP preparation.



Impact of CAP LAP accreditation on Laboratory Practice

A comparative analysis of laboratory practice before and after CAP LAP accreditation demonstrated statistically significant improvements across all domains (p-values <0.001) the improvement observed in self-inspection process, inspection, quality management system, staff training and competency, turnaround times, error reporting and corrective action processes, equipment maintenance and calibration, patient safety practice, staff engagement and morale, communication with clinicians, Laboratory safety program and data integrity and LIS policies. process shows a consistent pattern of increase in agreement and increases in strong agreement. For example, disagreement on checklist self-inspection fell from 18% to 0%, while strong agreement rose from 32% to 84%. In staff competency assessment, strong agreement increased from 36% to 86%. Turnaround time strong agreement grew from 43% to 78%, and error reporting strong agreement rose from 36% to 83%. Similarly, strong agreement on equipment maintenance rose from 38% to 84%, and patient safety prioritization from 42% to 88%. Staff morale and communication with clinicians also improved, with strong agreement reaching 77% and 76%, respectively Overall, these results confirm a consistent and significant improved in laboratory performance following the CAP LAP accreditation. The comparative analysis of laboratory practice before and after CAP LAP accreditation are summarized in Table 3.

Table 3: Comparative analysis of laboratory practice before and after CAP LAP accreditation.

Variable	Category	Before	After	Marginal homogeneity	
		N (%)	N (%)	p-value	Sig.
self-inspection process	**Strongly Disagree**	**18 (6%)**	**0 (0%)**	**<0.001**	**S**
Inspection process	Disagree	54 (18%)	0 (0%)		
	Neutral	72 (24%)	18 (6%)		
	Agree	60 (20%)	30 (10%)		
	Strongly Agree	96 (32%)	252 (84%)		
	Strongly Disagree	6 (2%)	0 (0%)	<0.001	S
	Disagree	60 (20%)	0 (0%)		
	Neutral	66 (22%)	12 (4%)		
	Agree	78 (26%)	48 (16%)		
	Strongly Agree	90 (30%)	240 (80%)		
Quality management system was well-established	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	36 (12%)	0 (0%)		
	Neutral	75 (25%)	6 (2%)		
	Agree	81 (27%)	51 (17%)		
	Strongly Agree	108 (36%)	243 (81%)		
Staff competency and training were regularly assessed	Strongly Disagree	6 (2%)	0 (0%)	<0.001	S
	Disagree	75 (25%)	0 (0%)		
	Neutral	36 (12%)	6 (2%)		
	Agree	75 (25%)	36 (12%)		
	Strongly Agree	108 (36%)	258 (86%)		
Turnaround times met clinical needs	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	27 (9%)	0 (0%)		
	Neutral	60 (20%)	18 (6%)		
	Agree	84 (28%)	48 (16%)		
	Strongly Agree	129 (43%)	234 (78%)		
Error reporting and corrective action processes were effective	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	39 (13%)	0 (0%)		
	Neutral	93 (31%)	6 (2%)		
	Agree	60 (20%)	45 (15%)		
	Strongly Agree	108 (36%)	249 (83%)		
Equipment maintenance and calibration were consistently performed	Strongly Disagree	6 (2%)	0 (0%)	<0.001	S
	Disagree	27 (9%)	0 (0%)		
	Neutral	78 (26%)	6 (2%)		
	Agree	75 (25%)	42 (14%)		
	Strongly Agree	114 (38%)	252 (84%)		
Patient safety was prioritized in daily operations	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	48 (16%)	0 (0%)		
	Neutral	66 (22%)	6 (2%)		
	Agree	60 (20%)	30 (10%)		
	Strongly Agree	126 (42%)	264 (88%)		
Staff engagement and morale were high	Strongly Disagree	18 (6%)	0 (0%)	<0.001	S
	Disagree	51 (17%)	0 (0%)		
	Neutral	63 (21%)	12 (4%)		
	Agree	72 (24%)	57 (19%)		
	Strongly Agree	96 (32%)	231 (77%)		

Variable	Category	Before	After	Marginal homogeneity	
		N (%)	N (%)	p-value	Sig.
Communication between lab and clinicians was effective	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	45 (15%)	0 (0%)		
	Neutral	69 (23%)	18 (6%)		
	Agree	90 (30%)	54 (18%)		
	Strongly Agree	96 (32%)	228 (76%)		
Laboratory safety program	Strongly Disagree	0 (0%)	0 (0%)	<0.001	S
	Disagree	60 (20%)	0 (0%)		
	Neutral	48 (16%)	6 (2%)		
	Agree	90 (30%)	42 (14%)		
	Strongly Agree	102 (34%)	252 (84%)		
Data integrity and LIS policies	Strongly Disagree	21 (7.14%)	0 (0%)	<0.001	S
	Disagree	57 (19.39%)	0 (0%)		
	Neutral	48 (16.33%)	6 (2%)		
	Agree	54 (18.37%)	42 (14%)		
	Strongly Agree	114 (38.78%)	252 (84%)		

Marginal homogeneity test showed statistically significant differences between pre- and post-accreditation responses ($p < 0.001$).

Patient Safety Culture

Patient safety culture ratings showed significant improvement after CAP LAP accreditation ($p < 0.001$). Before accreditation, the highest proportions were at level 4 (34.78%) and level 3 (30.43%), while only 21.74% rated it at the highest level (5) and 13.04% rated it as low (2). After accreditation, there was amarked shift, with 69.57% rating safety culture at level 5, 28.26% at level 4, and only 2.17% at level 2. This demonstrates a strong positive impact of CAP LAP accreditation on patient safety culture perceptions. The results illustrating the improvement in safety culture following CAP LAP accreditation are summarized in Table 4.

Table 4: Improvement in Patient Safety Culture Following CAP LAP Accreditation.

Variable	Rating	N (%)	Marginal homogeneity	
			p-value	Sig.
How would you rate the laboratory's patient safety culture before (CAP LAP)? (Linear scale 1–5)	2	36 (13.04%)	<0.001	S
	3	84 (30.43%)		
	4	96 (34.78%)		
	5	60 (21.74%)		
How would you rate the laboratory's patient safety culture after (CAP LAP)? (Linear scale 1–5)	2	6 (2.17%)		
	4	78 (28.26%)		
	5	192 (69.57%)		

Perceived improvements in laboratory performance and patient safety

Respondents reported strong improvements in several parameters of laboratory performance following CAP LAP accreditation. 72% reported significant increase in staff awareness of patient safety protocols, 50% reported significant reduction in patient laboratory error and 28% reported transformational reduction. 48% of respondents thought that communication between laboratory staff and clinicians was very effective, and 38% thought it was effective. 66% of respondents thought that corrective actions were very timely, and 22% thought they were timely. Regarding prevention of adverse outcomes, 50% reported significant improvement, 32% transformational, overall, 80% of respondents consider CAP LAP accreditation is extremely valuable for improving laboratory performance, the results are summarized in Table 5.

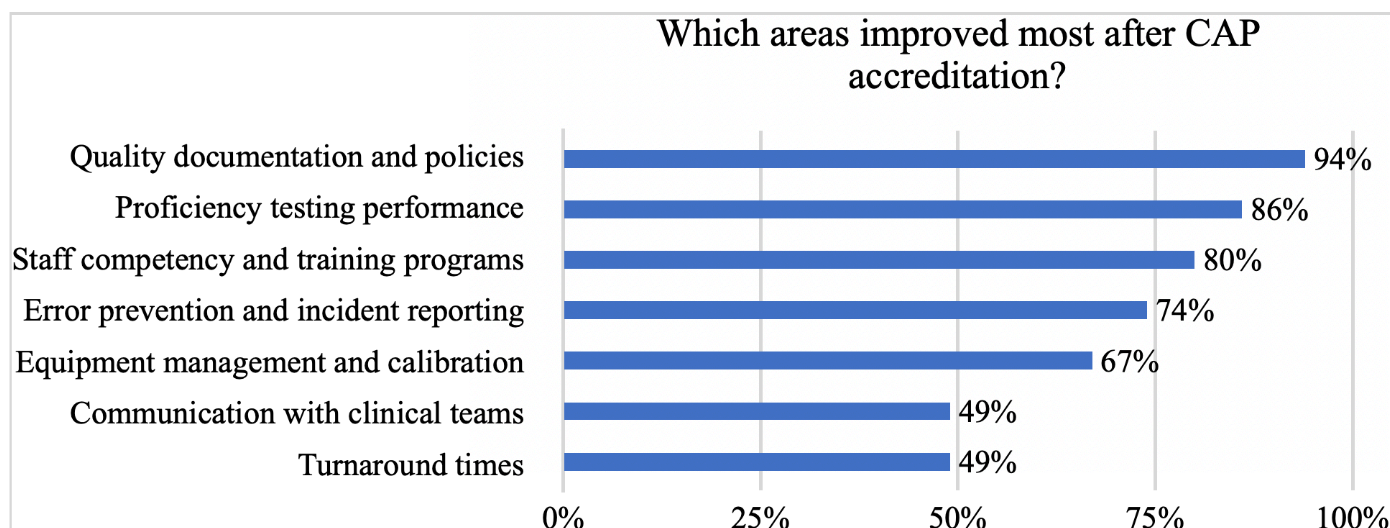
Table 5: Perceived improvements in laboratory performance and patient safety following (CAP LAP).

Variable	Category	(%)
Staff awareness of patient safety protocols has:	Decreased	6 (2%)
	Significantly increased	216 (72%)
	Slightly increased	54 (18%)
	N/A	24 (8%)
To what extent has (CAP LAP) reduced patient-related laboratory errors?	Moderate reduction	24 (8%)
	No reduction	6 (2%)
	Significant reduction	150 (50%)
	Slight reduction	12 (4%)
	Transformational reduction	84 (28%)
	N/A	24 (8%)
How effective is communication between laboratory staff and clinicians regarding patient safety after (CAP LAP)?	Effective	114 (38%)
	Neutral	18 (6%)
	Very effective	144 (48%)
	N/A	24 (8%)
How timely are corrective actions implemented after identifying safety issues?	Neutral	12 (4%)
	Timely	66 (22%)
	Very timely	198 (66%)
	N/A	24 (8%)
Has (CAP LAP) improved your lab's ability to prevent adverse patient outcomes?	Moderately	18 (6%)
	Significantly	150 (50%)
	Slightly	12 (4%)
	transformational	96 (32%)
	N/A	24 (8%)
How valuable do you consider (CAP LAP) for improving laboratory performance?	Extremely valuable	240 (80%)
	Moderately valuable	6 (2%)
	Valuable	54 (18%)

Respondents were asked to identify which operational areas within their laboratories showed the greatest improvement following CAP LAP accreditation.

The areas most improved after accreditation were quality documentation and policies (94%) and proficiency testing performance (86%). Staff competency and training programs also showed strong gains (80%), followed by error prevention and incident reporting (74%). Equipment management and calibration improved in 67% of laboratories. Meanwhile, turnaround times and communication with clinical teams were reported improved in 49% each. The results illustrate the domains where the impact of accreditation was most strongly perceived are summarized in Figure 2.

Figure 2: Areas improved after CAP LAP accreditation.



Discussion

The present study evaluates the perceived impact of college of American pathologist laboratory accreditation program CAP LAP on laboratory process and patient safety our come. Our findings indicate that accreditation was associated with improvements in numerous operational and quality-related areas of laboratory services, Participants reported greater adherence to standard operating procedures, improved documentation practices, and stronger implementation of quality management systems following accreditation.

Comparison with other studies

The demographic profile of our respondents is consistent with regional patterns observed in the Middle-East. Our findings that 54% Laboratories are accredited for more than 6 year and the distribution of the hospital-based lab 28% and Independent/ private 72% are aligned with Westgard, 2021 [22] found that 50.79% of the laboratories in the region are CAP accredited labs and there's an interesting mix between public and private hospitals in the Middle East. Our finding that the approximate number of staff in 52% of the laboratories from 20-50 and 10% is less than 20 employees is aligned with Westgard, 2021) [22] reported that there is a peak number of what we might call "mid-sized" modest laboratories in the region. The staff composition of 37% Laboratory Directors, 22% Technologists/Technicians,17% Quality Managers, 20% Section Heads/Supervisors supported by Parikh, & Rupani, (2025). [23] documenting those participants including Quality Managers, Head of laboratory, Technicians, and technical staff.

Our study findings that the main challenges during CAP LAP preparation were dominated by documentation requirements (68%) and (47%) training and competency gap supported by Alsayyah and Almulhim [1] reported that the documentation and training are primary concern during the preparation of thee accreditation, additionally our study findings where (66%) of the respondent reported Time management and 50% reported Cost and resource constraints are the main challenges during preparation of accreditation supported by AbdelWareth et al. [12] reported that time management and resources allocation are critical part. Our study finding where (48%) of the respondents reported that Staff resistance to change is one of the main challenges supported by Abdurabuh et al. (2024) [16] reported that accreditation process needs active change management.

The finding from the comparative study demonstrated that the CAP LAP accreditation has statistically signification impact on different area of laboratory practice. 12 variables assessed with ($p<0.001$ for each) provide evidence that there is improvement in medical laboratories performance following the CAP LAP accreditation. This aligns with Alsayyah and Almulhim [1] who found, showing that laboratories with more than accreditation more likely to have standards operating procedures ($P=0.039$) and review their policies more often ($P=0.002$) in Saudi Arabia. The huge increase in laboratories reporting well-established quality management systems (from 36% to 81% Strongly Agree) illustrates the foundational significance of certification in developing lasting quality culture this align with AbdelWareth et al. [12] who concluded that accreditation 'establishes a firm platform on which any new organization can construct a lasting quality improvement culture, this is further supported by Andiric et al. [21], who described how laboratory strengthening initiatives toward accreditation generate a 'culture of quality' and enable discovery of inadequacies. The observed improvement in patient safety prioritization (from 42% to 88% Strongly Agree) and error reporting effectiveness (36% to 83% Strongly Agree) following accreditation is aligns with Peter et al. [11], who documented that accreditation reduces testing errors and attendant decreases in inappropriate treatment. Abdurabuh et al. [16] also established a 'strong connection between accreditation and better patient safety' in Saudi healthcare facilities. The significant improvement in staff competency assessment (36% to 86% Strongly Agree) and staff engagement (32% to 77% Strongly Agree) after accreditation report one of the major pre-accreditation challenges identified in our study (staff resistance, 48%). This aligns with Hawkins' [17,18] observation that accreditation requirements drive laboratories to take greater responsibility for activities beyond analytical phases, including training and competency

development. The remarkable improvements in turnaround times (43% to 78% Strongly Agree), equipment maintenance (38% to 84% Strongly Agree), and communication with clinicians (32% to 76% Strongly Agree) reflect the CAP LAP accreditation benefits. These operational gains align with Alkhenizan and Shaw's [14], systematic review which found reliable evidence that accreditation programs improve both processes of care and clinical outcomes across healthcare system. The remarkable improvements in laboratory safety programs (from 34% to 84% strongly agreeing) and data integrity/LIS policies (from 38.78% to 84% strongly agreeing) demonstrate the comprehensive impact of CAP LAP accreditation. These areas, which are often unnoticed in routine laboratory operations, become systematized through the accreditation process, reducing risks to both patients and laboratory personnel.

The findings provide persuasive evidence that laboratory professionals perceive major improvements in patient safety and laboratory performance with CAP LAP accreditation The consistently high positive responses across all studied domains support the significance of accreditation as essential factor for quality advancement in medical laboratories. The finding that 72% of respondents indicated greatly enhanced awareness and 18% reported slightly increased awareness of patient safety practices following CAP LAP accreditation supports the program's success in fostering a safety-conscious culture. This aligns with Alsayyah and Almulhim's [1] survey of Saudi laboratory practitioners, which indicated that accreditation has had a positive impact on laboratory practices and patient laboratory testing in Saudi Arabia. This heightened awareness is important to the CAP accreditation philosophy, which demands laboratories to build comprehensive quality management systems that infuse safety awareness into daily operations. The perceived reduction in laboratory errors following accreditation is prominent: 50% of respondents reported significant reduction, 28% reported transformational reduction, and only 2% reported no reduction. These perceptions are strongly validated by the literature and aligns with Peter et al. [11] who reported that laboratory accreditation "has the potential to improve the quality of health care for patients through the reduction of testing errors and attendant decreases in inappropriate treatment". Similarly, Serteser et al. [15] demonstrated how ISO 15189 accreditation assures patient safety through robust quality management systems, noting that "applying the performance improvement strategies focusing on different phases in total testing process will significantly reduce the errors and therefore will improve the patient safety", also The systematic study by Alkhenizan and Shaw [14] gives the largest framework, stating that "accreditation programs improve the process of care provided by healthcare services" and "improve clinical outcomes of a wide spectrum of clinical conditions". Zima [13] further indicates that "accreditation of labs improves facilitation of accurate and rapid diagnostics, efficiency of treatment and reduction of errors in the laboratory process". The improvement in interdisciplinary communication following CAP LAP accreditation is supported, with 48% reporting very effective and 38% effective communication. This finding addresses a major determinant of patient safety that extends beyond the laboratory's walls. This aligns with Hawkins [17&18] who explored whether laboratory quality improvement through accreditation satisfies expectations and concluded that accreditation bodies require clear and effective procedures for patient/sample identification and communication of critical results. Hawkins [17&18] also stated that laboratories can play a greater role in minimizing errors and enhancing patient safety through better clinical interaction, The finding that 66% of respondents reported extremely timely and 22% reported timely implementation of corrective actions steps after recognizing safety issues illustrates the extensive quality management systems that CAP LAP accreditation establishes. This aligns with AbdelWareth et al. [12], who underlined that accredited laboratories build "a sustainable quality improvement culture" that permits methodical methods to quality management in their implementation evaluation of CAP and ISO 15189 accreditation. Andiric et al. [21] described how laboratory strengthening programs toward accreditation establish "a culture of quality in laboratories and allow the identification of gaps", allowing the institutionalization of timely remedial actions, The perception that CAP LAP accreditation significantly (50%) or transformationally (32%) improved laboratories' ability to prevent adverse patient outcomes. This aligns with Howanitz [20] who gave practical lessons for increasing patient safety through error reduction in laboratory medicine, similarly Milner and Holladay [2] claimed that laboratories serve as the "core for health systems building," stating that increasing laboratory quality through certification has rippling effects throughout healthcare systems. The finding that 80% of respondents consider CAP LAP extremely valuable and 18% consider it valuable for improving laboratory performance (98% total positive responses) provides strong confirmation from the laboratory professionals who experience accreditation personally. This aligns with Alsayyah and Almulhim's [1] observation that Saudi laboratory practitioners recognize the impact of accreditation in their daily practice. Similarly, AbdelWareth et al. [12] concluded their implementation evaluation by noting that "accreditation is a perfect means toward building quality medical laboratories in a diverse workforce environment and improving patient safety, also the analysis by Alkhenizan and Shaw [14] provides strong evidence that across many healthcare settings, healthcare professionals consistently value accreditation for its role in enhancing processes and outcomes

Limitation

First, the study utilized a cross-sectional design, which captures participants' impressions at a specific point in time and so cannot establish causal linkages between accreditation and observed improvements. In addition, the data were collected through self-administered questionnaires, depending on respondents' perceptions of laboratory practices before and after accreditation rather than on explicitly measurable performance indicators. Second, the study did not use objective laboratory information system (LIS) data, such as verified turnaround time records, laboratory error reports, near-miss registers, or quality indicator dashboards. As a result, the findings represent perceived improvements rather than statistically documented changes in laboratory performance or patient safety outcomes. Third, the sample distribution exhibited a regional imbalance, with the

majority of replies emanating from laboratories in Saudi Arabia. Furthermore, the study did not obtain data from laboratories that were not participating in the CAP Laboratory Accreditation Program CAP LAP, which limits comparisons between accredited and non-accredited laboratories. Finally, response bias may have influenced the results, as those with more pleasant experiences with accreditation may have been more willing to engage in the survey. In addition, retrospective recall of procedures before accreditation may have led some respondents to overstate observed improvements.

Conclusion

In conclusion, this study highlights the significant role of the college of College of American Pathologists Laboratory Accreditation Program CAP LAP in improving laboratory processes and patient safety. This improvement was observed from the results of the study across all laboratory measured areas. CAP LAP not only standardizes procedures but also develops organizational culture, increases staff participation, and promotes continual quality improvement. The study indicates considerable improvements across laboratory practice areas, with strong evidence supporting the impact of CAP LAP accreditation and a clear relationship between enhanced safety culture and fewer laboratory errors. It is recommended in future studies to include the non-CAP or ISO 15189-accredited laboratories in future research; this could enable comparative impact analysis and further validate the unique contribution of CAP LAP to laboratory excellence and patient safety.

Patents

The authors declare that no patents are associated with this work.

Supplementary Materials

No supplementary materials are available for this manuscript

Author Contributions

Abdelaziz Sanad contributed to the study design, data collection, and initial data analysis. He also assisted in drafting and revising the manuscript, contributed to algorithm development, computational analysis, and technical validation of the study.

Mai S. Mabrouk led the study concept and design, oversaw the computational framework implementation, ensured technical rigor and reproducibility, contributed to data interpretation and manuscript revision. provided guidance on methodology, assisted with validation of the computational framework, and critically reviewed the manuscript for scientific accuracy

Samir Y. Marzouk participated in methodology development, supervised the data analysis, and contributed to interpreting the results.

Funding

This research received no external funding.

AI Statement

No artificial intelligence tools were used in the design, analysis, or writing of this manuscript.

Institutional Review Board Statement

The survey conducted in this instance was voluntary, anonymous, and non-interventional and exempt from IRB review

Informed Consent Statement

Informed consent was obtained electronically from all participants prior to survey submission.

Data Availability

The anonymized survey dataset supporting the findings of this study is available from the corresponding author upon reasonable request. Researchers interested in accessing the dataset should contact A. Sanad at asanad83@yahoo.com

Acknowledgments

The artificial intelligence tool (ChatGPT) was used solely for language editing and Grammer check.

Conflicts of Interest

The authors declare no conflicts of interest.

Abbreviations

CAP	College of American Pathologists
CLIA	Clinical Laboratory Improvement Amendments
CLSI	Clinical and Laboratory Standards Institute
CMS	Centers for Medicare & Medicaid Services
FDA	Food and Drug Administration
ISO	International Organization for Standardization
ISQua	The International Society for Quality in Health Care
LAP	Laboratory Accreditation Program
TAT	Turnaround Time
LIS	Laboratory information system
NS	Non-significant
OSHA	Occupational Safety and Health Administration.
QC	Quality Control
QMS	Quality management system
S	Significant
SOPs	Standard operating procedures
SPSS 27	Statistical Package for Social Science

References

1. Alsayyah, A.; Almulhim, A. What Has Been the Impact of Laboratory Accreditation in Saudi Arabia? A Cross-Sectional Survey of Laboratory Practitioners. J. Datta Meghe Inst. Med. Sci. Univ. 2023;18,398–404. https://doi.org/10.4103/jdmimsu.jdmimsu_80_23.
2. Milner, D.A., Jr.; Holladay, E.B. Laboratories as the Core for Health Systems Building. Clin. Lab. Med. 2018;38,1–9.
3. International Organization for Standardization (ISO). ISO 15189:2012- Medical Laboratories- Requirements for Quality and Competence; ISO: Geneva, Switzerland, 2012.
4. Joint Commission International (JCI). Laboratory Accreditation Program. Available online:https://www.jointcommissioninternational.org/what-we-offer/accreditation/accreditation-programs/laboratory/ (accessed on 15 October 2025).
5. CBAHI. Saudi Central Board for Accreditation of Healthcare Institutions - Portal. Available online: https://portal.cbahi.gov.sa/ (accessed on 15 October 2025).
6. GAHAR. General Authority for Healthcare Accreditation and Regulation (CARE Standards). Available online: https://gahar.gov.eg/ (accessed on 15 October 2025).
7. Merrick, T. (Ed.). College of American Pathologists Commission on Laboratory Accreditation: Laboratory Accreditation Manual; College of American Pathologists: Northfield, IL, USA, 1996.
8. Hamlin, W.B. Requirements for Accreditation by the College of American Pathologists Laboratory Accreditation Program. Arch. Pathol. Lab. Med. 1999;123,465–467.
9. Shahbaz, A.; Younas, M.; Khanzada, F.A.; Meer, K.R.; Ahmad, L.; Ameer, R.; Zahid, M.A. Analyzing the Impact of College of American Pathologists Laboratory (CAP) Accreditation Program on Post-Analytical Error. Res. Med. Sci. Rev. 2024;2,586. Available online: https://thermsr.com (accessed on 29 November 2025).
10. Mark, H.F.L.; Watson, M.S. Evolving Standards of Practice for Clinical Cytogenetics. R. I. Med. J. 1994;77,375–377.
11. Peter, T.F.; Rotz, P.D.; Blair, D.H.; Khine, A.A.; Freeman, R.R.; Murtagh, M.M. Impact of Laboratory Accreditation on Patient Care and the Health System. Am. J. Clin. Pathol. 2010;134, 550–555. https://doi.org/10.1309/AJCPH1SKQ1HNWGHF.
12. AbdelWareth, L.O.; Pallinalakam, F.; Ibrahim, F.; Anderson, P.; Liaqat, M.; Palmer, B.; Mirza, I. Fast Track to Accreditation: An Implementation Review of College of American Pathologists and ISO 15189 Accreditation. Arch. Pathol. Lab. Med. 2018, 142, 1047–1053.
13. Zima, T. Accreditation of Medical Laboratories- System, Process, Benefits for Labs. J. Med. Biochem. 2017;36, 231–237. https://doi.org/10.1515/jomb-2017-0023.
14. Alkhenizan, A.; Shaw, C. Impact of Accreditation on the Quality of Healthcare Services: A Systematic Review of the Literature. Ann. Saudi Med. 2011;31,407–416. https://doi.org/10.4103/0256-4947.83204.
15. Serteser, M.; Coskun, A.; Inal, T.C.; Unsal, I. How ISO 15189 Laboratory Accreditation Assures Patient Safety. J. Med. Biochem. 2012;31,345–352.
16. Abdurabuh, A.; Hamid, M.D.; Che Hassan, C.R.; Fatani, M.I. Evaluating the Impact of Hospital Accreditation on Patient Safety Culture in Saudi Arabia Healthcare Facilities. J. Multidiscip. Healthc. 2024, [ahead of print], 5021–5033.
17. Hawkins, R. Laboratory Quality Improvement through Accreditation: Are We Meeting Expectations? Clin. Chem. Lab. Med. 2012;50,1115–1122. https://doi.org/10.1515/cclm-2011-0944.
18. Hawkins, R. Managing the Pre- and Post-Analytical Phases of the Total Testing Process. Ann. Lab. Med. 2012;32,5–16. https://doi.org/10.3343/alm.2012.32.1.5
19. College of American Pathologists (CAP). Laboratory Accreditation Program. Available online:https://www.cap.org/laboratory-improvement/accreditation/accreditation-program/laboratory-accreditation-program (accessed on 29 November 2025).
20. Howanitz, P.J. Errors in Laboratory Medicine: Practical Lessons to Improve Patient Safety. Arch. Pathol. Lab. Med. 2005;129,1252–1261.
21. Andiric, L.R.; Chavez, L.A.; Johnson, M.; Landgraf, K.; Milner, D.A. Strengthening Laboratory Management toward Accreditation: A Model Program for Pathology Laboratory Improvement. Clin. Lab. Med. 2018;38,131–140.
22. Westgard, S. (2021, November). The 2021 Middle-East QC survey results. Westgard QC. https://westgard.com/qc-applications/basic-qc-practices/2021-middle-east-qc-survey-results.html
23. Parikh KD, Rupani MP. Exploring the impact of national laboratory accreditation on quality and practices: a qualitative study from a government medical college in western India. BMC Health Services Research. 2025;25(1):1005. doi:10.1186/s12913-025-13195-6.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review Article

A review of the methods available for the detection of antibodies against transglutaminase and deaminated gliadin in Celiac Disease, traditional and emerging technologies.

Diana Landoni[1,2], Gerson Dierley Keppeke[3,4] *

[1]Escuela de Graduados, Facultad de Medicina, Universidad de la República, Montevideo, Uruguay
[2]Laboratorio de Análisis Clínicos, LAC, Montevideo, Uruguay
[3]Departamento de Ciencias Biomédicas, Facultad de Medicina, Universidad Católica del Norte, Coquimbo, Chile
[4]Disciplina de Reumatologia, Departamento de Medicina, Universidade Federal de São Paulo, Brasil

Article Info

Corresponding Author:

Gerson Dierley Keppeke

Av. Larrondo 1281, Departamento de Ciencias Biomédicas
Universidad Católica del Norte, Coquimbo, Chile

ORCID: 0000-0003-0660-2857

E-mail: gerson.keppeke@ucn.cl

Keywords

Celiac Disease, Transglutaminase, Deamidated Gliadin; Laboratory testing, Autoantibodies

Abstract

Celiac Disease (CD) is an immune-mediated enteropathy where serological testing, centered on IgA anti Tissue Transglutaminase 2 (anti-TG2) and IgA anti Endomysium antibodies (anti-EMA), is the diagnostic cornerstone. Anti-TG2 is recognized as the most sensitive marker, while anti-EMA is the most specific. This review examines the performance characteristics, advantages, and limitations of traditional and emerging technologies for CD serology.

Traditional platforms include ELISA - a cost-effective, open platform used for quantitative anti-TG2 and anti-DGP detection and IIF for anti-EMA, which, despite being labor-intensive and operator-dependent, maintains high specificity by detecting anti-TG2 in its native conformation. Newer technologies prioritize automation in view of clinical laboratory demands.

CLIA and FEIA are fully automated systems that offer enhanced analytical sensitivity and rapid workflow. Meta-analyses confirm that the diagnostic accuracy of IgA anti-TG2 is statistically equivalent across CLIA (Sens. 0.98, Spec. 0.97), FEIA (Sens. 0.97, Spec. 0.99), and ELISA (Sens. 0.96, Spec. 0.97). However, CLIA often requires higher diagnostic thresholds for biopsy-sparing protocols. Emerging solutions include Multiplex Flow Immunoassays (MFI) /Microarrays, which enable simultaneous multi-isotype antibody detection with high concordance to reference assays (PPA 96.0 %, NPA 98.0 % for IgA anti-TG2) while Point-of-Care Tests (POCTs) provide rapid, equipment-free screening for primary care, though they exhibit moderate sensitivity compared to laboratory assays.

The overall trend favors automated and multiplexed methods, improving efficiency and supporting biopsy-sparing strategies based on robust quantitative autoantibody profiling.

Introduction

Celiac Disease is not an "allergy" to gluten, but an autoimmune disease

Celiac disease (CD) is a chronic, immune-mediated enteropathy that predominantly affect the small intestine of genetically predisposed individuals, precipitated by the exposure to food that contains gluten [1]. This disease affects females more frequently than males in a 3:1 ratio, as in most autoimmune diseases (AID), with first-degree relatives being at a higher risk of developing CD than the general population [2,3]. The genetic predisposition depends on the presence of HLA-DQ2/DQ8 (necessary condition, but not sufficient to develop CD on its own). Approximately 95.0 % of people with CD possess the HLA-DQ2 gene, while the remaining patients often carry HLA-DQ8 [4]. The worldwide prevalence of CD varies but is estimated to be approximately 1 % of the population [1].

Transglutaminase 2 (TG2) belongs to a family of enzymes that catalyze post-translational protein modifications through cross-linking or the incorporation of primary amines [5]. These reactions generate products that are highly resistant to mechanical stress and proteolytic degradation, explaining their prevalence in tissues where structural stability is essential [6]. Of the eight transglutaminases identified in the genome, six have been characterized through purification or recombinant methods [7].

Tissue transglutaminase 2 (TG2) participates in processes such as cell death, differentiation, matrix stabilization, and cell adhesion. It is best known for catalyzing covalent cross-links between proteins via transamidation, forming stable bonds between glutamine and lysine residues. This activity enhances extracellular matrix (ECM) stability and promotes cell attachment following enzyme externalization [6]. In addition to cross-linking, TG2 can catalyze deamidation, converting glutamine residues into glutamic acid. Given its broad tissue and cellular distribution, and its multiple enzymatic and non-enzymatic functions, the transglutaminase family is implicated in the pathogenesis of several diseases, including CD. In CD, TG2 is overexpressed and localized to the enterocyte brush border, cytoplasm, and ECM [8]. In patients with HLA predisposition, gluten is recognized as toxic and triggers inflammatory responses involving both the innate and adaptive pathways of immunity. Although the exact mechanism by which different antigens become targets of the immune system is not completely understood, it is thought that in CD the T cells activated by gluten cross-react with TG2 through epitope spreading [2]. The TG2 enzyme deamidates gluten peptides such as gliadin, making them better available to bind to MHC II on the surface of dendritic cells, activating the immune response [1] and increasing gliadin immunogenicity [5]. This process results in the activation of B lymphocytes that differentiate into plasma cells producing autoantibodies against TG2 (among others), leading to a full autoimmune response. Autoantibodies against TG2 (anti-TG2) contribute to CD pathogenesis by interfering with tissue repair, promoting inflammation and increasing permeability through the cells that form the outer lining of the small intestine [6]. The autoantibodies bind to TG2 and inhibit its enzymatic function, which in turn disrupts the tight junctions between the epithelial cells leading toincreased permeability or "leaky gut". In this condition pathogens filter through the cells undetected and undeterred by the immune system towards the lamina propria. Once bound to TG2 they form immune complexes that activate the immune system via the complement cascade and enhance pro-inflammatory cytokine production, mainly through IL-15, driving intraepithelial lymphocyte activation and epithelial damage. These autoantibodies also provoke a dysregulation of the pathways that remodel the extracellular matrix, impairing angiogenesis and contributing to fibrosis by exacerbating mucosal injury and perpetuating inflammation and tissue damage [7].

Gliadin is the main protein contained in wheat and due to its aminoacidic sequence and structure it is difficult to digest by human endopeptidases [8]. The proline/glutamine-rich fragment (peptide p31-43) is resistant to digestion by gastric, pancreatic, and intestinal proteases and it plays a crucial role in epithelial stress, increased permeability, and innate immune activation. The undigested peptide residues can have different biological activities such as inducing an innate or adaptive immune response related to their ability to bind to HLA-DQ 2.5, DQ 8 or DQ 2 [9,10].

Antibodies against deamidated gliadin are specific to modified gliadin peptides that are deamidated by TG2 as discussed above. These posttranslational modifications are believed to contribute to CD by augmenting gliadin immunogenicity [11]. Beyond the increased immunogenicity of TG2-mediated deaminated peptides, gliadin itself can trigger inflammation in a context of increased intestinal permeability. Gliadin is a plant component, thus acts as a foreign antigen and can directly stimulate innate immune cells, even in individuals without CD. Once in contact with gluten the intestinal epithelium releases zonulin which increases permeability and the translocation of gliadin to the lamina propria [2]. Overall, CD pathogenesis involves gluten-induced activation of adaptive immunity driven by TG2-deamidated peptides that stimulate Th1 $CD4^+$ T cells, IFN-γ production, and a specific autoantibody response leading to intestinal damage [12].

The cornerstone of CD diagnosis is the laboratory testing of autoantibodies.

In addition to anti-TG2, in CD one can find IgA and IgG -an-bodies against gliadin (anti-gliadin antibodies - AGA) as well as against deamidated gliadin peptides (anti-DGP) [13], (Box 1). AGA can be found in other enteropathies or even in healthy individuals, which lowers its specificity and may lead to false positive results. For this reason, other autoantibodies are preferred, such as anti-TG2, anti-DGP or anti-EMA [1]. The endomysium is a layer of connective tissue that surrounds muscle fibers and thus is filled with TG2 enzymes either in the ECM or on the cell surface, allowing for the autoantibodies to be able to detect the enzyme in its natural conformational structure within the tissue. Essentially, the anti-EMA assay detects the presence of anti-TG2 in the sample, although other autoantibodies could be present - such as antibodies against other transglutaminases TG3 and TG6 [14]. These antibodies have been related to extraintestinal manifestations of CD [7] such as dermatitis herpetiformis (DH), and neurological complications like ataxia [15].

Many AIDs are linked, meaning the presence of one disorder seems to increase the risk of developing others. CD is no exception and is linked to selective IgA deficiency (sIgAD), type 1 diabetes, autoimmune thyroid disease and some chromosomal abnormalities [16].
The case of sIgAD is especially relevant, since IgA plays a homeostatic role regulating the intestinal mucosa. Patients that lack the proper concentrations of this antibody are more at risk of developing an immune reaction to certain foreign antigens, in the case of CD, to gluten. IgA levels increase gradually after birth until they reach peak levels during adolescence (with normal levels between 61-365 mg/dl) [17].
The cornerstone of CD diagnosis is the laboratory testing of autoantibodies, mainly IgA anti-TG2 and IgA anti-EMA, combined with total serum IgA in order to detect those cases that also present sIgAD [1,17,18]. Anti-DGP (IgA and IgG) and IgG anti-TG2 are also available, the former being considered useful in the diagnosis of CD in infants up to two years of age who have not yet fully developed IgA antibodies [12,18,19]. These types of serological tests have become important due to the fact that they are very convenient to perform (high throughput and automatization) and have lower risks than the duodenal biopsy [20]. Anti-TG2 have also been found to identify adults with mucosal lesions in their duodenum accurately helping clinicians avoid unnecessary invasive procedures [21]. The new 2025 Updated Guidelines on the Diagnosis and Management of Celiac Disease in Adults proposed by the European Society for the Study of Celiac Disease (ESsCD), recommends that routine use of IgA anti-EMA serology is no longer necessary, as long as validated high-performance IgA anti□TG2 assays are used [22]. Due to the labor-intensive and limited availability, IgA anti-EMA could be reserved for unclear cases to ensure diagnostic accuracy [22].

Anti-TG2 are considered pathogenic in CD, meaning they are drivers of the disease and keep perpetuating tissue damage. However, AGA and anti-DGP are not autoantibodies since the human genome does not contain the gene for gliadin (Box 1). This distinction between autoantibodies (anti-TG2 and anti-EMA) and regular antibodies (AGA and anti-DGP) is important because they reflect on different biological and immunological processes, the former are evidence of loss of tolerance by the immune system while the latter are showing a response towards external dietary antigens. As such, AGA and anti-DGP offer more insight into how the body reacts to gliadin when it is not digested properly.
All in all, anti-TG2 and anti-EMA are the primary diagnostic tools for CD, while anti-DGP is used when anti-TG2 results are uncertain, especially in young children or in patients with SIgAD. Usually, a combination of these tests along with a histological evaluation of the small intestine (via biopsy) and a correct anamnesis is the best way to diagnose CD [23]. The standard and most effective way to manage CD is following a gluten free diet (GFD) [24,25]. Considering that the digestion-resistant gliadin peptide p31-43 plays a crucial role in epithelial stress, increased permeability, and innate immune activation, avoiding it will essentially eliminate or considerably put a stop to CD major clinical complications, mainly involvingthe progression of mucosal injury, and perpetuation of inflammation and tissue damage. Follow up of these patients is not always simple. Studies have shown that IgG antibodies (for anti-TG2, anti-DGP and anti-EMA) persist longer than IgA antibodies, which makes sense considering their sources - long versus short-lived plasma cells - and the different roles the antibody isotypes play. IgA antibodies are more predominant in the intestinal mucosa and titers drop rapidly after GFD adherence, while IgG antibodies may show a longer-term immune response, especially if related to extra-intestinal symptoms [18]. In the American Gastroenterology Association clinical practice update of 2019 experts concluded that serology was more useful for follow-up in the case of children than in adults, and that lowering IgA anti-TG2 could help identify those patients with a healing intestinal mucosa [16].

Objective

In this manuscript, we review the methods available for detection of autoantibodies in celiac disease, including anti-transglutaminase and anti-deaminated gliadin, both traditional and emerging technologies.

Box 1: Are AGA and anti-DGP really autoantibodies? The correct answer is no.
We asked different online platforms with “Artificial Intelligence”, made of large language models, what they “thought” autoantibodies were. Responses were along the lines of autoantibodies are antibodies that mistakenly target and attack the body’s own tissues or organs. They are produced by the immune system when it loses the ability to distinguish between “self” and “non-self”, known as loss of immune tolerance, this is a hallmark of autoimmune diseases.
An expert in the field would further argue that autoantibodies actually represent a spectrum within an individual’s *antibodyomics*, including the physiologically important natural autoantibodies, which help with the immune system´s function of eliminating damaged cells and dead-cell remains, as well as other autoantibodies that are found in high titer in a significant proportion of people but are not related to any common autoimmune disease, such as the famous anti-DFS70, discussed elsewhere [26].

Technologies for detection of antibodies, biomarkers in celiac disease

Methodologically speaking, this study is not a systematic review but rather a narrative review. Extensive search was applied to identify studies related to different technologies and platforms for the detection of antibodies in CD, focused on anti-TG2, anti-DGP and anti-EMA. Still, the flowchart for the screening of studies following PRISMA-S method is presented in Supplementary Figure 1.

Traditional Technologies

ELISA (Enzyme-Linked Immunosorbent Assay)

ELISA is the most commonly used method for detecting anti-TG2 and anti-DGP. Based on the principle of antigen–antibody binding, antigens used to coat the wells may be recombinant or native, these can be extracted from sources such as animal tissues, wheat, or other biological materials.

Historically, different generations of ELISA tests have been defined based on the origin of the substrates, with early tests using TG2 derived from guinea pig liver before human derived or recombinant TG2 demonstrated superior performance, Box 2 details the case for DGP.

Dieterich *et al.* identified TG2 as the autoantigen against which the detection of anti-reticulin antibodies (ARA) and anti-EMA tests were made in 1997, ushering in a new era for CD serology testing with the development of specific ELISA testing [27]. This technique is highly sensitive, easy to perform, automatize and can be done at relatively low cost since it doesn't require highly trained personnel or dedicated equipment. However different commercial kits (which usually include all the different reagents and solutions necessary to perform these tests) can have varying specificities, heavily influenced by the source of their antigens. For clinical diagnostics the assays that should be used are those with CE-IVD or FDA approval, or that have been approved by country-specific regulatory agencies. One should not assume that just because the platform is similar (ELISA), assays from different manufacturers will have similar performance as our group has extensively discussed in previous publications [28]. This holds true for any antibody detection kit, production lots may vary within the same brand, kits may vary between brands, and everything varies depending on the quality of the antigen used to coat the wells of each kit. These variabilities may produce false positives or false negatives [14].

Antibody detection can be modulated by modifying either the substrate or the secondary anti-human immunoglobulin (Ig) antibody. Using TG2 as a substrate enables the detection of anti-TG2 antibodies, while coating wells with DGP allows for the detection of anti-DGP antibodies. Similarly, modifying the secondary antibody with specificity for different Ig isotypes allows for the identification of IgG, IgA, or both [29].

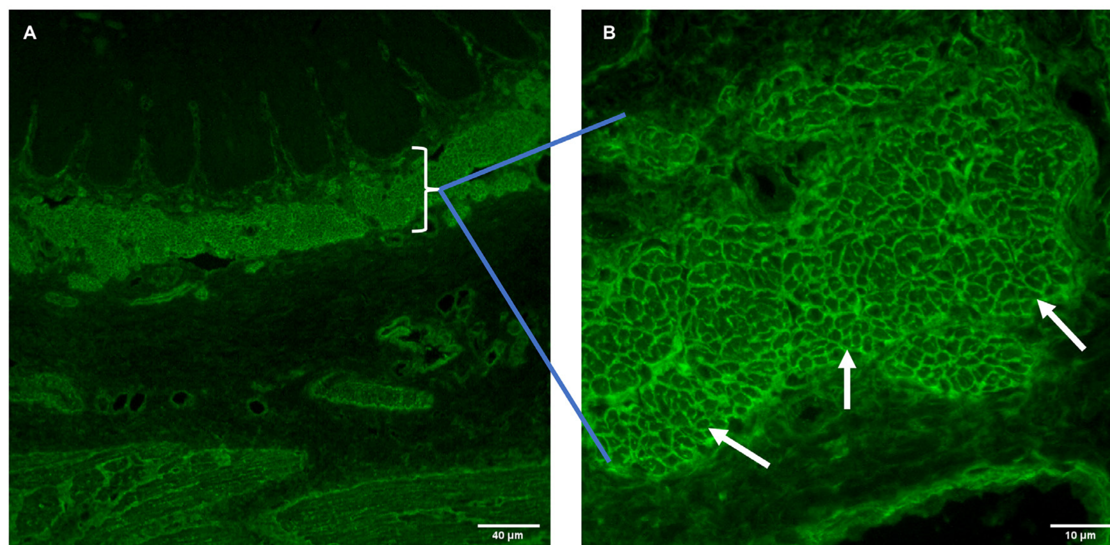
Box 2: Obtaining deaminated gliadin for immunoassays.
DGP peptides used in solid-phase immunoassays for anti-DGP detection are generated through a multi-step process that begins with the extraction of gliadin from wheat flour. Gliadin - a prolamin protein fraction of gluten - is isolated when using solvent extraction methods, these typically employ ethanol or aqueous alcohol solutions, which separate gliadin from other gluten components such as glutenins [30].

Following extraction, the critical step of deamidation is performed enzymatically by using purified recombinant TG2, which mimics the physiopathological modifications present in patients with CD and ensures the preservation of the conformational epitopes recognized by anti-DGP. This reaction occurs under physiological simulated pH and temperature conditions, with TG2 preferentially targeting glutamine residues in motifs such as QXP or QXXF(Y/W/M/L/I/V), which are essential for HLA-DQ2/DQ8 binding in the pathogenesis of CD [6,7,30].

Indirect Immunofluorescence (IIF) Used for detecting anti-EMA.

The traditional substrate used for detecting anti-EMA by IIF is monkey esophagus, however there are several other alternative substrates such as human umbilical cord, human esophagus or primate liver. Several authors from the 1990s published articles trying to discourage the use of endangered species for organ harvesting to be used as substrate, however all of these current options have problems either with availability or with ethical and environmental concerns [29,31,32]. Volta *et al.* compared results for IgA anti-EMA on monkey esophagus versus human umbilical cord for patients with CD under the premise that the human umbilical cord is: 1. commonly available, 2. rich in reticulin fibers (endomysium) that surround smooth muscle fibers, and 3. that unlike other human tissues it does not contain IgA, therefore limiting cross reactivity [31]. No differences were found between doing the test in either substrate in this study (meaning a similar performance) either with the CD patients or the control groups. This test can be labor intensive if it is done manually. Several semi-automated machines are available to help laboratories run the immunodetection part of the assay. The glass slide is coated with the substrate in several "wells" and on each well the machine pipettes a dilution of the patient's sera. After the slides are ready, the technician can observe the characteristic honeycomb-like immunofluorescence pattern in the microscope (Figure 1). The interpretation is subjective to the technicians' training and skill. High background noise and non-specific staining can hinder readers from confidently emitting a result. As with all IIF techniques, anti-EMA testing has its limitations due to tissue autofluorescence, reagent instability and differences between reagent brands and even kit lots. It is important to highlight that patients must be consuming gluten when this test is performed for diagnosis.

Figure 1: Characteristic honeycomb-like pattern in samples with anti-TG2 antibodies.



Indirect immunofluorescence assay with monkey Esophagus as substrate (Kit NOVA Lite, Ref. 704155, Inova Diagnostics). A representative sample with anti-TG2 antibodies, positive for both IgG and IgA, was applied and the assay was performed following manufacturer recommendations. (A) 10x objective. (B) 40x objective. The characteristic honeycomb-like pattern (arrows) can be observed with 400x magnification

The results of this test can be reported as either qualitative (positive or negative) or in a semi-quantitative form by reporting titers from double dilutions. Despite differences in initial screening dilutions, Murray *et al.* found that across several laboratories in the USA IgA anti-EMA testing usually reports high sensitivities (60.0-100 %) and specificities (80.0-100 %), however both of these parameters are affected by the initial screening dilution [29,33–36]. Even though the concordance between laboratories was high (kappa coefficient: 0.739), the laboratory that used the highest screening dilution (1:20) had a statistically lower sensitivity than the rest [33]. Anti-EMA values are usually linked to higher degrees of histological lesions and villous atrophy and its titers also change once the patient has implemented a GFD [24,29,37].

The main difference between an IgA and IgG anti-EMA is the FITC conjugate immunoglobulin used, if one is testing IgG then an immunoglobulin against the Fc portion of an IgG antibody should be used. When using monkey esophagus as a substrate the pattern observed is the same. Nevertheless, as we have stated before, it is recommended to use this test exclusively in the case of sIgAD as is the case for any other IgG antibody used for diagnosis in this disease [38].

Anti-EMAs are not the only identified targets for CD-related autoantibodies, others have been described as well, but their use has dwindled in the past decades since ease of use and automatization of the more commonly used tests increased. Other examples of targets are: calreticulin, zonulin, desmin, reticulin and antibodies against jejunum [39].

Anti-reticulin antibodies (ARA) were originally described in 1977 in patients with DH, CD and Crohn's disease in rat stomach, kidney, and liver tissue (SKL) where the R1 pattern (peritubular staining in the kidney or periportal tissue in the liver) was associated with CD [40]. Both IgA and IgG ARA can be detected [29] and seem to be closely related to anti-EMA. Hällström *et al.* found positive IgA antibodies for both ARA and anti-EMA in 91.0 % of an untreated adult CD patient cohort, while 83.0 % of a cohort of DH patients had both autoantibodies [40]. Even so, the study could not confidently differentiate both diseases (CD and DH) with these (or any other) tests. Both types of antibodies were found to decrease with a GFD. Although both autoantibodies seem to be closely related, they do not appear to have the same antigen specificity. Mäki described these tests as rodent-type reticulin (reticulin) and primate-type reticulin (endomysium) in the early 90s [39]. In 2003 Koroponay-Szabó *et al.* reported a study with TG2 knockout mouse tissue in which they conclude that diagnostically relevant anti-EMA, ARA and jejunal antibody (JEA) found in patients with CD and DH were clearly and exclusively anti-TG2 autoantibody dependent [41]. More recent papers have declared ARA obsolete due to its failure to measure up to other tests used for routine diagnostics [42]. Rodent substrate offers a lower sensitivity and larger inherent subjectivity due to the fact that the R1 pattern in SKL must be detected in the three different rodent tissues. With the later identification of TG2 as the target antigen for anti-EMA, and that anti-DGP proved to be a more specific CD biomarker, and their subsequent test developments, the algorithm for CD diagnosis changed thus, ARAs fell further in disuse [27,43]. As Korponay-Szabó and collaborators state in their 2003 publication, ARA and anti-EMA are not antigen dependent or even species dependent but rather detect the same phenomena in different tissue types, as both detect anti-TG2 autoantibodies [41]. These autoantibodies that have not endured the test of time have not demonstrated to be S3MA2RT2, as Fritzler *et al.* described in 2021 [44]. They were either not Specific, Sensible or Scalable, Measurable using conventional technologies, did not *Add value* to clinical management or were not Actionable or Realistic, or were neither Titrable nor addressed the Temporal Timing of the disease, and so fell into the "death-valley" of autoantibodies [44].

Caja et al. describe how anti-EMA is a superior test for anti-TG2 ELISA due to the fact that it shows a greater correlation to the presence of HLA DQ2 or DQ8, which is not always the case for individuals who are anti-TG2 seropositive [14]. This, and because anti-EMA is a "more stable test" makes it hold its place as the "gold-standard" serology test, although the concept of gold-standard changes over time. A 2006 review by Lewis and Scott discusses how anti-EMA more often has a higher specificity while tests for anti-TG2 antibody have a higher sensitivity [45]. Thus, the majority of the current guidelines such as the 2020 European Society for Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) and the 2023 American College of Gastroenterology (ACG) guideline, recommend performing solid-phase assays such as anti-TG2 and anti-DGP as screening and confirming antibody positivity with the more specific anti-EMA assay. Yet, as previously mentioned, some 2025 updated guidelines, as from ESsCD, recommends the decrease of routine IgA anti-EMA in adult CD diagnostics [22,34,46–48].

Recent wide-spread Technologies

Chemiluminescence Immunoassay (CLIA): Provides enhanced diagnostic accuracy with better automation and sensitivity.

CLIA applies the same antigen–antibody principles as ELISA but detects immune complexes through light emission expressed as relative luminescence units (RLU). The antigen, usually human recombinant TG2 is immobilized on paramagnetic or glass microbeads This technique requires specialized equipment to be able to run the test and detect the light emitted. Compared with conventional colorimetric ELISA or fluorescence based FEIA (fluorescence enzyme immunoassay) (described ahead), CLIA offers enhanced analytical sensitivity, broader dynamic range, and full automation. Paramagnetic-bead chemistry provides a larger reactive surface and lower background noise, improving precision and reducing inter-assay variability.

A systematic review and meta-analysis found pooled sensitivity 0.98 (95.0 % CI 0.95–0.99) and specificity 0.97 (95.0 % CI 0.94–0.99) for IgA anti-TG2 by CLIA, values comparable or slightly superior to ELISA and FEIA [49]. The analysis also noted delayed normalization of anti-TG2 titers during GFD follow-up when measured by CLIA, an important consideration for long term monitoring. Elsewhere, in a large biopsy-correlated series using a CLIA platform for IgA anti-TG2 (QUANTA Flash by Werfen), sensitivity and specificity reached 98.2 % and 98.4 %, respectively [50]. The optimal biopsy-saving thresholds were ≈ 15 × ULN (upper limit of normal) in adults (350 CU) and ≈ 28 × ULN in children (560 CU), higher than the 10 × ULN rule adopted for ELISA assays [23]. This data supports CLIA's diagnostic reliability but indicates that method-specific cut-offs must be applied.

In summary, CLIA provides standardized, quantitative detection of anti-TG2 and anti-DGP antibodies with high automation and throughput, though it requires dedicated instrumentation and calibration traceability [51].

Fluorescence Enzyme Immunoassay (FEIA): Offers quicker response times and higher throughput than ELISA.

FEIA is a variant of ELISA, the difference being that instead of a colorimetric reaction at the end of the test, the secondary antibody is attached to a fluorescent molecule (usually FITC), thus fluorescence intensity is detected. This change enables the use of different detection methods, improving sensitivity and allowing for the measurement of lower concentration of antibodies. FEIA could be considered as an evolution of the ELISA principle, offering shorter turnaround times and higher analytical throughput. The use of fluorescence-based detection also facilitates multiplexing capabilities and broader dynamic range, contributing to improved quantification and reproducibility. Commercial FEIA platforms (e.g., Phadia EliA systems) are fully automated, enabling continuous sample loading and standardized calibration protocols. Both CLIA and FEIA surpass ELISA in speed, automation, and traceability, but CLIA generally achieves slightly higher sensitivity at very low antibody titers. The main limitation in both is higher costs since they are closed systems with dedicated equipment.

According to the meta-analysis by Pjetraj et al., FEIA for IgA anti-TG2 reached pooled sensitivity of 0.97 and specificity of 0.99, comparable to CLIA (0.98 / 0.97) and ELISA (0.96 / 0.97) [49]. Despite this, FEIA showed faster response times and better reproducibility in multicenter evaluations. These results confirm FEIA as a robust, high-performance alternative suitable for both screening and follow-up of CD.

Emerging technologies

Multiplex Flow Immunoassay (MFI) and microarray: Emerging as an alternative with greater sensitivity for different antibody types.

Multiplex assays enable the detection of multiple autoantibodies in a single test and can be classified regarding the ability that they have for specific autoantibody identification [28,52]. For example, if one had a test for detecting IgA and IgG anti-DGP simultaneously (as in the case for INOVA CLIA anti-DGP screen test) one could not identify if a positive sample was positive for antibody isotype IgA, IgG, or both. In the case of a Line Blot, one could identify every individual autoantibody through the antigen adsorbed in a line on the nitrocellulose membrane that would appear in the case of sample positivity.

More modern formats for MFIs are bead-based (microspheres with unique fluorescent signatures covered with the antigens such as BioPlex 2200 platform [53]), or microarray-based technologies that use "spots" of antigens adsorbed on a nitrocellulose membrane, polymer, or hydrogel. These assays have been developed thanks to the use of molecular biology cloning and micro-printing technologies and have been adapted to large-scale laboratories through technology that allows for these assays to run automatically and with a higher throughput [28]. The conservation of the selected proteins in their native conformation is essential for optimal autoantibody detection. The test principle is the same as with ELISA, CLIA and IIF; involving the formation of immune complexes between the substrate antigen, patient autoantibody and the secondary detection antibody, which according to the preferred detection method, can be conjugated to a fluorophore or chemiluminescent dye.

A recent example of this technological evolution is the MosaiQ AiPlex Celiac Disease Microarray Solution, an IVDR-CE–marked microarray system designed for the MosaiQ platform.

A analytical comparison conducted in a reference laboratory in Wales, UK evaluated the MosaiQ AiPlex CD microarray from AliveDx - a fully automated, single-use multiplex immunoassay - for the simultaneous semi-quantitative detection of IgA and IgG anti-TG2 and anti-DGP antibodies [54]. When compared with routine ELISA, FEIA, and CLIA methods (EliA Celikey, QUANTA Flash, QUANTA Lite, and ORGENTEC assays), the AiPlex system demonstrated a positive percent agreement (PPA) of 96.0

% and negative percent agreement (NPA) of 98.0 % for IgA anti-TG2, and PPA 78.0 % / NPA 93.0 % for IgG anti-TG2. Overall concordance with standard assays was high (90.0 % positive and 95.0 % negative agreement), confirming the analytical equivalence and reliability of this microarray platform for CD serology. This platform is presented as a robust and efficient alternative for high-throughput laboratories, capable of simultaneously analyzing multiple antibody isotypes while minimizing sample handling and operator variability [54].

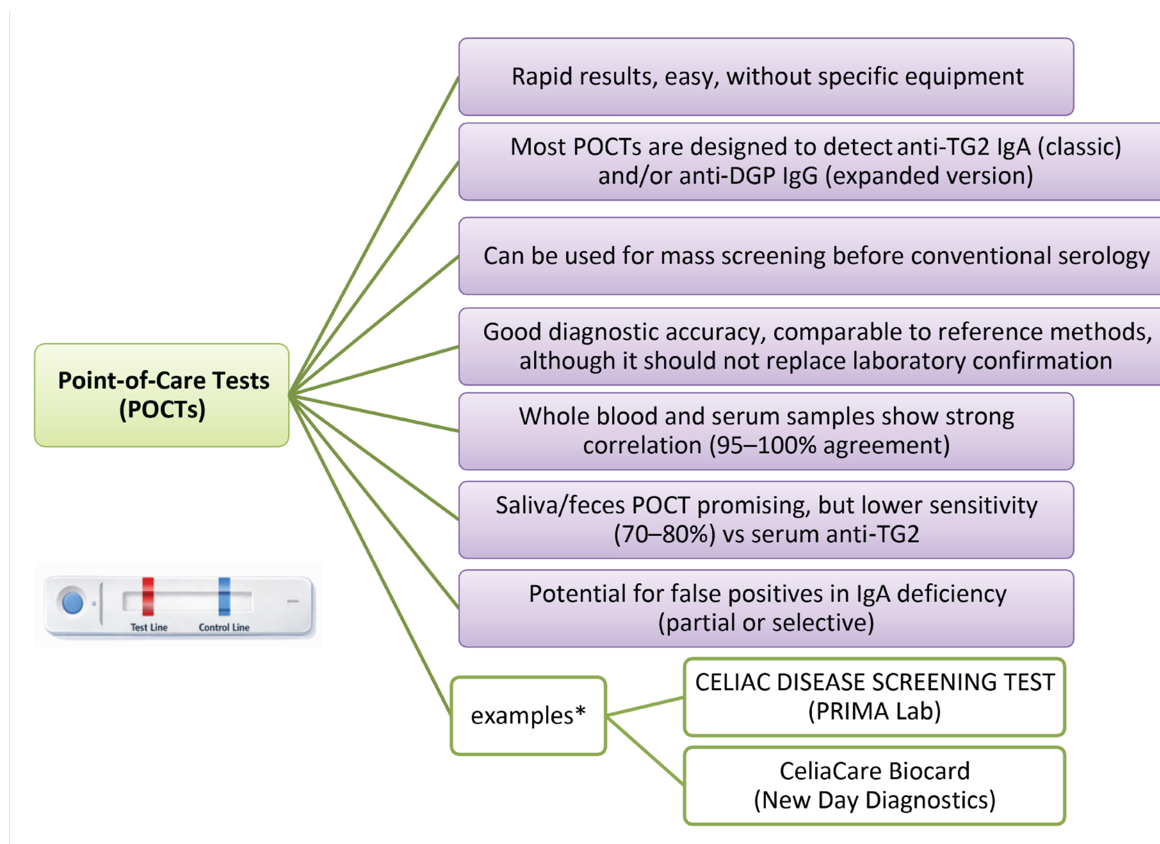
Point-of-care tests (POCT): Rapid tests developed for use in settings with limited resources or in a general physician's office. POCTs offer immediate results with moderate sensitivity with no need of laboratory equipment or expert personnel [29,55], such as the CeliaCare Biocard (from New Day Diagnostics) and the CELIAC DISEASE SCREENING TEST (from PRIMA Lab) among others. These are single-use immunochromatographic or dot-blot assays in which patient antibodies bind immobilized recombinant TG2, forming visible immune complexes detected without instrumentation [56].

Baldas et al. first described a new dot-blot assay specifically designed for ambulatory settings based on the recognition of human, recombinant TG2 and compared it to IIF (anti-EMA) and ELISA methods [57]. The test could be performed in 20 minutes and presented a sensitivity of 100 % and a specificity of 98.0 %, for which they concluded it was highly accurate in detecting untreated patients (Figure 2). This test is able to detect patients with sIgAD due to concomitant detection of IgA and IgG antibodies, has a low cost and can be performed with minimal handling. Another advantage is that it can be performed with either serum or whole blood drop [58], although in both cases it is diluted and not used directly. Nemec et al. used lateral flow tests and reported that when they used the whole blood drop test the sensitivity was slightly lower, 96.0 % than in the test that required serum (100 %) in comparison to ELISA tests [58]. They were unable to identify the patients that had sIgAD since that commercial brand of POCT could not measure IgG anti-TG2.

This methodology was tested in both adults [59] and children [60] and has improved since Baldas's description in the 2000s. In the study by Korponay-Szabó et al. [60] 2690 children of 6 years of age were tested with a 5-minute immunochromatographic card detecting IgA anti-TG2 antibodies in whole blood. In this study the test yielded 78.1 % sensitivity and 100 % specificity, while sensitivity decreased to 65.1 % when compared to combined laboratory testing for IgA and IgG anti-EMAs by IIF. Alarida et al. also reported a low sensitivity in their study performed on Libyan children in a cohort of low pre-test probability [61] concluding that rapid determination of IgA or IgG anti-TG2 on whole blood drop was not the test of choice for screening purposes.

One of the advantages of this kind of testing is that it could be useful for population screening, especially in countries that have routine mandatory screening for children. However, there are still concerns regarding the cost-effectiveness of these measures as public policy, and regarding the adequate follow up and continued care for these families and patients [55]. Other issues are that POCT may have quality control concerns when handled by users with little experience, not to mention that rapid results are rarely decisive in CD [16]. POCTs should not replace laboratory confirmation in CD (Figure 2).

Figure 2: Summary of the main advantages and limitations of POCTs for CD.



*It is illustrating its operational simplicity and clinical applicability compared with conventional laboratory-based serological methods. *Two examples of rapid lateral-flow or dot-blot assays are mentioned in the Figure; however, numerous commercial POCTs are available from different manufacturers, such as Simtomax (Augurix Diagnostics), Simple CD1WB (Operon), MyTest Celiac (Cooper Consumer Health) and Xeliac Pro (EUROSPITAL SpA), to mention a few examples. While some POCTs are well studied and clinically validated, others are consumer-oriented kits with limited or no independent clinical validation.*

Strengths and weaknesses of different platforms

Various publications go over comparisons between different techniques. However, before going into detail, it is important to remember that even if one compares different techniques and platforms (CLIA vs IIF, IIF vs ELISA, etc.) one should always consider that there can be significant variability within the same technique among different brands, especially for ELISAs, as discussed above.

A recent systematic review by Pjetraj *et al.* evaluated the diagnostic accuracy of IgA anti-TG2 assays across different technologies, including CLIA, ELISA, and FEIA [49]. Pooled results demonstrated excellent overall performance for all methods, with CLIA sensitivity 0.98 (95.0 % CI 0.95–0.99) and specificity 0.97 (95.0 % CI 0.94–0.99), FEIA 0.97/0.99, and ELISA 0.96/0.97. No statistically significant differences in diagnostic accuracy were found between the three assay types, indicating that each provides reliable detection of IgA anti-TG2 antibodies when validated and properly calibrated. The authors also noted minor inter-method variability in antibody-titer normalization during gluten-free diet follow-up, underscoring the need for harmonized cut-off definitions and traceable calibration standards (Table 1). Taken together, this evidence supports the analytical equivalence between modern immunoassay platforms and reinforces their collective value in the accurate, non-invasive serologic diagnosis of CD.

For the diagnosis of CD, anti-TG2 IgA represents the most sensitive marker and the basis of current screening algorithms, while IgA anti-EMA remains the most specific assay, closely correlating with the presence of villous atrophy. Anti-DGP antibodies, particularly of the IgG isotype, serves as useful complementary markers in younger children or in patients with sIgAD. Comparative studies and meta-analyses have quantified the relative diagnostic performance of these antibodies, providing the basis for modern serologic algorithms and biopsy-sparing criteria.

A study by Volta *et al.* (2023) mentions the high variability in accuracy and tracks diagnostic precision among several serologic markers for CD such as IgA anti-TG2, IgA anti-EMA and both IgA and IgG for anti-DGP [62]. With a total of 5.098 pediatric and adult patients among all 44 original studies, they found that IgA anti-TG2 had a slightly higher sensitivity than IgA anti-EMA (93.4 % vs 92.8 %) and IgG/IgA anti-DGP had lower sensitivities (81.8 % and 83.8 % respectively). The highest specificity was as expected for IgA anti-EMA (99.0 %), followed by IgG anti-DGP (96.4 %), IgA anti-TG2 (95.8 %) and IgA anti-DGP (92.1 %) [59,62].

Lewis and Scott (2006) compared anti-EMA and anti-TG2 finding very high sensitivities (93.0 %) and specificities (more than 99.0 % and 98.0 % respectively) for both assays in the context of patients with CD that presented villous atrophy [45].

They also reviewed the difference between using human recombinant TG2 vs guinea pig TG2 (human recombinant tests had high sensitivities and specificities with a 95.0 % CI) as well as the difference between using monkey esophagus vs human umbilical cord for anti-EMA studies (they were both very similar, 93.1 % vs 92.9 % in sensitivities and 99.1 % vs 99.7 % specificities, with a 95.0 % CI) [45].
Table 1 summarizes the characteristics, advantages, and limitations of the main serological technologies for CD. Traditional methods such as ELISA and IIF remain the most widely used, with ELISA predominating as an open, versatile, and cost-effective platform adaptable to different reagents and instruments. IIF has exceptional specificity for anti-EMA antibodies but remains labor-intensive and operator-dependent,meaning higher costs as well. In contrast, automated methods such as CLIA and FEIA offer enhanced analytical sensitivity, broader dynamic range, and random-access automation, aligning with modern high-throughput laboratory workflows. The advent of multiplex and microarray systems has enabled simultaneous multi-isotype antibody detection within a single reaction, combining efficiency with excellent analytical concordance to reference assays. POCTs extend diagnostic accessibility to low-resource or primary-care settings, providing rapid, equipment-free results suitable for preliminary screening (Figure 2). Collectively, these technologies illustrate a clear transition from traditional manual assays toward integrated, automated, and multiplexed solutions that improve diagnostic precision, workflow efficiency, and patient accessibility.

Table 1: Comparison of Traditional and Emerging Technologies.

Technology	Characteristics	Strengths	Platform weaknesses and critical points of the experimental procedures	Marker	Diagnostic Performance*		References
					Sensitivity Range	Specificity Range	
ELISA	Quantitative solid-phase assay for detecting anti-TG2 and anti-DGP antibodies. The most widely used platform for CD serology.	- Open platform, compatible with multiple reagents and instruments. - Cost-effective, simple, and well validated. - Extensively standardized and accepted by regulatory agencies.	- Requires batch testing and manual handling. - Variable performance across manufacturers and antigen sources. - Limited to single-analyte detection. - Must be careful with standard curve calibration and definition of cut-offs.	IgA anti-TG2	90.0% - 98.0%	95.0% - 99.0%	[1,23,36,62,63]
				IgG anti-TG2	~70% or lower in IgA-competent	94.0% - 100%	[1,18,22]
				IgG anti-DGP	80.0% - 97.0%	86.7% - 99.0%	[1,7,64]
Indirect Immunofluorescence (anti-EMA test)	Detects anti-EMA antibodies using tissue substrates (monkey esophagus or human umbilical cord).	- High specificity (“gold standard” for anti-EMA). - Good confirmatory test. - Detects antibodies against TG2 in their native conformational state.	- Subjective pattern interpretation requiring skilled personnel. - Semi-quantitative. - Requires fluorescence microscopy. - Labor-intensive; ethical and supply issues with animal substrates.	IgA anti-EMA	83.8% - 98.0%	97.5% - 100%	[1,22,23,62,65,66]
CLIA (Chemiluminescence Immunoassay)	Fully automated, quantitative detection based on light emission from acridinium- or isoluminol-labeled complexes.	- High analytical sensitivity and dynamic range. - Rapid, random-access workflow and excellent reproducibility.	- Closed system, higher instrument cost and potential procurement issues in some countries, vendor lock-in. - Requires calibration traceability and maintenance. - High dynamic range (0.1-4965 CU); requires method-specific cut-offs.	IgA anti-TG2	95.3% - 98.3%	94.0% - 99.0%	[49,50,67]

Technology	Characteristics	Strengths	Platform weaknesses and critical points of the experimental procedures	Marker	Diagnostic Performance*		References
					Sensitivity Range	Specificity Range	
FEIA (Fluorescence Enzyme Immunoassay)	Variant of ELISA using fluorescent detection.	- Faster turnaround and higher throughput than ELISA. - Broad dynamic range and multiplexing potential.	- Closed automated system with higher per-test cost and potential procurement issues in some countries. - Limited flexibility for antigen/cut-off customization. - Slightly lower sensitivity than CLIA at very low titers.	IgA anti-TG2	~97%	98.0% - 99.0%	[21,49,62]
Multiplex Flow Immunoassay / Microarray	Detects multiple autoantibodies (anti-TG2, anti-DGP) simultaneously on bead- or microarray-based platforms.	- Enables multiplexing of IgA/IgG isotypes in a single test. - May even incorporate internal IgA-deficiency check. - High throughput, minimal sample volume, and automation. - Comparable accuracy to reference methods (PPA 96.0%, NPA 98.0% for IgA anti-TG2).	- High setup cost and need for complex panels and robust assay standardization. - Limited clinical availability. - Interpretive software required.	Combined (anti-TG2/ DGP)	77.0% - 96.0%	98.0% - 100%	[53,68,69]
POCT (Point-of-Care Tests)	Lateral-flow or dot-blot rapid assays detecting anti-TG2 IgA/IgG +/- anti-DGP IgG.	- Rapid, low-cost, and equipment-free. - Suitable for field use and primary-care screening. - Whole-blood and serum results strongly correlated (95.0-100% agreement).	- Moderate sensitivity compared with laboratory assays. - Not reliable to rule out CD in high-suspicion clinical settings. - Potential false results in IgA deficiency. - Positive results should be confirmed with standard laboratory testing. - Limited quality control in untrained hands. - Many consumer-oriented kits with little-to-no clinical validation.	IgA/IgG anti-TG2	70.0% - 98.0%	90.0% - 98.0%	[56,65,67,70]

**Diagnostic performance of CD serological platforms in comparison with reference standards such as duodenal histology and/or gold-standard antibody assays.*

Conclusion

Advances in serologic technologies over the past decade - and how those are applied - have improved considerably, resulting in significant changes benefiting the patients such as allowing the diagnosis of CD without duodenal biopsies in defined scenarios. In children, the ESPGHAN 2020 pathway allows a biopsy-sparing diagnosis when IgA anti-TG2 is ≥10× ULN, confirmed on a second sample with anti-EMA and made within a shared decision-making process [64,71]. Large prospective validation studies have confirmed the safety of this approach: Wolf *et al.* demonstrated that children with IgA anti-TG2 ≥10× ULN plus anti-EMA positivity can be diagnosed without biopsy, achieving a PPV ≈ 99.0 %; Chokkalla *et al.* reported recently (2024) comparable accuracy for the biopsy-free approach versus the traditional method in North American pediatric cohorts [72,73].

Analytical innovations have also clarified assay-specific cut-offs and limitations. Previtali *et al.* (2018) showed that CLIA may require higher diagnostic thresholds than ELISA, while Pjetraj *et al.* (2024) confirmed that CLIA achieves sensitivity 0.98 and specificity 0.97 - comparable to conventional platforms [49,50]. Multiplex and multi-analyte systems also follow the same logic: in 2025 Gambino *et al.* validated a multiparametric IgA blot combining several gliadin and TG-related epitopes with 100 % specificity when ≥ 6 antigen markers were positive; the same year Zingone *et al.* showed that concurrent positivity for IgA anti-TG2 and IgG anti-DGP above 10×ULN predicted Marsh 3 histology in adults with 100 % PPV [68,69].

As is the case for many AIDs, assay standardization - in particular for autoantibody measurement - remains a major challenge, and CD is no exception. The lack of international standardization for antibody concentrations means that results obtained across different platforms are not directly comparable and often require local validation to support clinical decisions. A significant step toward addressing this gap is the recent international effort to harmonize CD serology through certified reference materials. In 2025, the Joint Research Centre of the European Commission, in collaboration with the IFCC, released the first certified reference material for anti–TG2 antibodies, both IgA and IgG (ERM-DA487_IFCC). This initiative provides a metrologically traceable calibration anchor, representing an important advance toward more reliable clinical cut-offs and standardized diagnostic pathways in CD.

Viewed together, these developments point towards a future in which most CD diagnoses will be biopsy-free, achieved through the integration of quantitative autoantibody profiles, HLA-DQ2/DQ8 genotyping, and emerging T-cell or mucosal-immune biomarkers [74]. Machine-learning and multi-omics tools, such as the XGBoost-based prescreening model of Dreyfuss *et al.*, demonstrate how artificial intelligence can recognize latent serologic or biochemical signatures before clinical diagnosis [75,76]. In this precision-diagnostic framework, intestinal biopsy will likely remain a problem-solving procedure, reserved for seronegative, equivocal, or refractory presentations rather than routine.

Another recent development, showcasing a future trend for laboratory testing in CD, is related to the new 2025 ESsCDupdated guidelines that further consolidate IgA anti-TG2 at the center of CD diagnosis in adults, recommending that routine anti-EMA is not required when high-performance anti-TG2 assays are used, reserving anti-EMA for equivocal cases [22]. However, we must emphasize that in the current guidelines from ESPGHAN and ACG, the recommendation is to perform solid-phase assays such as anti-TG2 and anti-DGP as screening and confirm positivity with the more specific anti-EMA assays [34,46,47].

List of Abbreviations

ACG: American College of Gastroenterology
AID: Autoimmune Disease
AGA: Anti-gliadin Antibodies
ARA: Anti-reticulin Antibodies
CD: Celiac Disease
CE-IVD: Conformité Européenne – In Vitro Diagnostic
CLIA: Chemiluminescence Immunoassay
CU: Conventional Units
DGP: Deamidated Gliadin Peptides
DH: Dermatitis Herpetiformis
ECM: Extracellular Matrix
ELISA: Enzyme-Linked Immunosorbent Assay
EMA: Anti-endomysium Antibodies
ESPGHAN: European Society for Paediatric Gastroenterology, Hepatology and Nutrition
ESsCD: European Society for the Study of Coeliac Disease
FDA: Food and Drug Administration
FEIA: Fluorescence Enzyme Immunoassay
GFD: Gluten-Free Diet
HLA: Human Leukocyte Antigen
IIF: Indirect Immunofluorescence
JEA: Jejunal Antibody
MFI: Multiplex Flow Immunoassay
NPA: Negative Percent Agreement
NPV: Negative Predictive Value
POCT: Point-of-Care Testing
PPA: Positive Percent Agreement
PPV: Positive Predictive Value
sIgAD: Selective IgA Deficiency
SKL: Stomach, Kidney, and Liver Tissue
TG2: Transglutaminase 2
ULN: Upper Limit of Normal
USA: United States of America

Declarations

Declaration of Conflict of interests

The authors declare that they have no conflicts of interest in relation to this manuscript.

Ethical Approval

Not applicable.

CRediT Author Statement

DL: Conceptualization, Methodology, Investigation, Writing – Original Draft.
GDK: Conceptualization, Methodology, Supervision, Writing – Review & Editing.

Funding
During development of this work, DL was supported by *Comisión Académica de Posgrado* (CAP) from Universidad de la República (UdelaR), Uruguay. GDK was supported by FAPESP (*Fundação de Amparo à Pesquisa do Estado de São Paulo*), Brazil. GDK was also granted a *Fondo de Fortalecimiento a la Investigación de Iniciación y Avanzada 2025*,
N°186/2025-VRIDT-UCN, Universidad Católica del Norte, Chile.

Data Availability Statement
Not applicable.

Acknowledgments

Disclosure on the use of AI tools
The manuscript was originally written in Spanish and Portuguese, the mother language of the authors, and translated to English using online AI tools such as ChatGTP (by OpenAI) and Gemini (by Google/Alphabet). The same tools were also applied to improve the structure and clarity of parts of the text. Figure 2 was partially generated using NotebookLM and ChatGTP, followed by manual adjustments. The online platforms with "Artificial Intelligence" mentioned in Box 1 for the question about autoantibodies were ChatGTP, Gemini and Elicit (https://elicit.com/). Last, as detailed in Supplementary Figure 1, identification and filtering of entries and references were carried out with the help of automation tools, such as Google's Gemini, ChatGPT, Zotero, among others, in other to build the list of studies that went for manual screening to be part of the reference list, following PRISMA-S method.

References

1. Bai JC, Ciacci C. World Gastroenterology Organisation Global Guidelines: Celiac Disease February 2017. J Clin Gastroenterol. 2017;51:755–768.https://doi.org/10.1097/MCG.0000000000000919
2. Kivity S, Ehrenfeld M. Can we explain the higher prevalence of autoimmune disease in women? Expert Rev Clin Immunol. 2010;6:691–694. https://doi.org/10.1586/eci.10.60
3. Singh P, Arora S, Lal S, Strand TA, Makharia GK. Risk of Celiac Disease in the First- and Second-Degree Relatives of Patients With Celiac Disease: A Systematic Review and Meta-Analysis. Am J Gastroenterol. 2015;110:1539–1548. https://doi.org/10.1038/ajg.2015.296
4. Cecilio LA, Bonatto MW. The prevalence of HLA DQ2 and DQ8 in patients with Celiac Disease, in family and in general population. ABCD Arq Bras Cir Dig São Paulo. 2015;28:183–185. https://doi.org/10.1590/S0102-67202015000300009
5. Martucciello S, Sposito S, Esposito C, Paolella G, Caputo I. Interplay between Type 2 Transglutaminase (TG2), Gliadin Peptide 31-43 and Anti-TG2 Antibodies in Celiac Disease. Int J Mol Sci. 2020;21:3673. https://doi.org/10.3390/ijms21103673
6. Griffin M, Casadio R, Bergamini CM. Transglutaminases: nature's biological glues. Biochem J. 2002;368:377–396. https://doi.org/10.1042/BJ20021234
7. Paolella G, Sposito S, Romanelli AM, Caputo I. Type 2 Transglutaminase in Coeliac Disease: A Key Player in Pathogenesis, Diagnosis and Therapy. Int J Mol Sci. 2022;23:7513. https://doi.org/10.3390/ijms23147513
8. Barone MV, Salvatore A. Pro-Inflammatory Nutrient: Focus on Gliadin and Celiac Disease. Int J Mol Sci. 2022;23:5577. https://doi.org/10.3390/ijms23105577
9. Herrera MG, Dodero VI. Gliadin proteolytical resistant peptides: the interplay between structure and self-assembly in gluten-related disorders. Biophys Rev. 2021;13:1147–1154. https://doi.org/10.1007/s12551-021-00856-z
10. Falcigno L, Calvanese L, Conte M, Nanayakkara M, Barone MV, D'Auria G. Structural Perspective of Gliadin Peptides Active in Celiac Disease. Int J Mol Sci. 2020;21:9301. https://doi.org/10.3390/ijms21239301
11. Diós Á, Srinivasan B, Gyimesi J, Werkstetter K, Valenta R, Koletzko S, et al. Changes in Non-Deamidated versus Deamidated Epitope Targeting and Disease Prediction during the Antibody Response to Gliadin and Transglutaminase of Infants at Risk for Celiac Disease. Int J Mol Sci. 2022;23:2498. https://doi.org/10.3390/ijms23052498
12. Volta U, Fabbri A, Parisi C, Piscaglia M, Caio G, Tovoli F, et al. Old and new serological tests for celiac disease screening. Expert Rev Gastroenterol Hepatol. 2010;4:31–35. https://doi.org/10.1586/egh.09.66
13. Bakhtiari S, Rostami-Nejad M. Chapter Three - Emerging markers in celiac disease. Makowski GS, editor. Adv Clin Chem. Elsevier Academic Press; 2025. Amsterdam. Vol 129. p. 123–89. https://doi.org/10.1016/bs.acc.2025.06.010
14. Caja S, Mäki M, Kaukinen K, Lindfors K. Antibodies in celiac disease: implications beyond diagnostics. Cell Mol Immunol. 2011;8:103–109. https://doi.org/10.1038/cmi.2010.65
15. Croall ID, Hadjivassiliou M, Sanders DS, Teh K, Biancardi AM, Trott N, et al. Persisting Transglutaminase 6 Antibodies in Neurological Gluten-Related Disorders. Ann Neurol. 2025; https://doi.org/10.1002/ana.78020
16. Husby S, Murray JA, Katzka DA. AGA Clinical Practice Update on Diagnosis and Monitoring of Celiac Disease-Changing Utility of Serology and Histologic Measures: Expert Review. Gastroenterology. 2019;156:885–889. https://doi.org/10.1053/j.gastro.2018.12.010
17. Poddighe D, Capittini C. The Role of HLA in the Association between IgA Deficiency and Celiac Disease. Birková A, editor. Dis Markers. 2021;2021:1–8.https://doi.org/10.1155/2021/8632861
18. Di Tola M, Bizzaro N, Gaudio M, Maida C, Villalta D, Alessio MG, et al. Diagnosing and Monitoring Celiac Patients with Selective IgA Deficiency: Still an Open Issue. Dig Dis Sci. 2021;66:3234–3241. https://doi.org/10.1007/s10620-021-07204-x
19. Bufler P, Heilig G, Ossiander G, Freudenberg F, Grote V, Koletzko S. Diagnostic performance of three serologic tests in childhood celiac disease. Z Für Gastroenterol. 2015;53:108–114. https://doi.org/10.1055/s-0034-1385704
20. Shiha MG, Nandi N, Raju SA, Wild G, Cross SS, Singh P, et

al. Accuracy of the No-Biopsy Approach for the Diagnosis of Celiac Disease in Adults: A Systematic Review and Meta-Analysis. Gastroenterology. 2024;166:620–630. https://doi.org/10.1053/j.gastro.2023.12.023

21. Iversen MN, Stribolt K, Hvas CL, Dige A. Diagnostic accuracy of IgA anti-tissue transglutaminase for the diagnosis of coeliac disease. Dan Med J. 2025;72:A03250187. https://doi.org/10.61409/A03250187
22. Al-Toma A, Zingone F, Branchi F, Schiepatti A, Malamut G, Canova C, et al. European Society for the Study of Coeliac Disease 2025 Updated Guidelines on the Diagnosis and Management of Coeliac Disease in Adults. Part 1: Diagnostic Approach. United Eur Gastroenterol J. 2025;00:1–32. https://doi.org/10.1002/ueg2.70119
23. Wang Z, Wu Q. Recent advances in the diagnosis of celiac disease. World J Gastroenterol. 2025;31:112388. https://doi.org/10.3748/wjg.v31.i37.112388
24. Tarar ZI, Zafar MU, Farooq U, Basar O, Tahan V, Daglilar E. The Progression of Celiac Disease, Diagnostic Modalities, and Treatment Options. J Investig Med High Impact Case Rep. 2021;9:23247096211053702. https://doi.org/10.1177/23247096211053702
25. Caio G, Volta U, Sapone A, Leffler DA, De Giorgio R, Catassi C, et al. Celiac disease: a comprehensive current review. BMC Med. 2019;17:142. https://doi.org/10.1186/s12916-019-1380-z
26. Ochs RL, Mahler M, Basu A, Rios-Colon L, Sanchez TW, Andrade LE, et al. The significance of autoantibodies to DFS70/LEDGFp75 in health and disease: integrating basic science with clinical understanding. Clin Exp Med. 2016;16:273–293. https://doi.org/10.1007/s10238-015-0367-0
27. Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, et al. Identification of tissue transglutaminase as the autoantigen of celiac disease. Nat Med. 1997;3:797–801. https://doi.org/10.1038/nm0797-797
28. Keppeke GD, Agustinelli RDA, Landoni DM, Cruvinel WM, Chan EKL, Andrade LEC. Chapter 88 - Future Direction of Laboratory Testing for Autoimmune Diseases. In: Schmitz JL, Detrick B, O'Gorman MRG, editors. Man Mol Clin Lab Immunol. 9th Ed. Wiley & ASM Press; 2024. Washington, D.C. p. 968–982. https://doi.org/10.1002/9781683674023.ch88
29. Barakauskas VE, Lam GY, Estey MP. Digesting all the options: Laboratory testing for celiac disease. Crit Rev Clin Lab Sci. 2014;51:358–378. https://doi.org/10.3109/10408363.2014.958813
30. Lexhaller B, Ludwig C, Scherf KA. Identification of Isopeptides Between Human Tissue Transglutaminase and Wheat, Rye, and Barley Gluten Peptides. Sci Rep. 2020;10:7426. https://doi.org/10.1038/s41598-020-64143-9
31. Volta U, Molinaro N, De Franceschi L, Fratangelo D, Bianchi FB. IgA anti-endomysial antibodies on human umbilical cord tissue for celiac disease screening: Save both money and monkeys. Dig Dis Sci. 1995;40:1902–1905. https://doi.org/10.1007/BF02208653
32. Izzi L. Anti-endomysium antibodies detection in the celiac disease screening: Indirect immunofluorescence pattern using umbilical cord sections as substrate. ALTEX (ISSN 1868-8551). 1998;15:141–143.
33. Murray JA, Herlein J, Mitros F, Goeken JA. Serologic Testing for Celiac Disease in the United States: Results of a Multilaboratory Comparison Study. Clin Diagn Lab Immunol. 2000;7:584–587. https://doi.org/10.1128/CDLI.7.4.584-587.2000
34. Rodríguez N, Buzzi C, González A, López V, Olano C, Ballesté R. Consenso uruguayo en el diagnóstico serológico de la enfermedad celíaca del adulto. Rev Médica Urug. 2024;40:1–22. https://doi.org/10.29193/RMU.40.3.s2
35. Cabo Del Riego JM, Núñez Iglesias MJ, García-Plata González C, Paz Carreira J, Álvarez Fernández T, Dorado Díaz A, et al. Evaluation of Celiac Disease by Minimally Invasive Biomarkers in a Spanish Pediatric Population. Int J Environ Res Public Health. 2022;19:5020. https://doi.org/10.3390/ijerph19095020
36. Infantino M, Merone M, Manfredi M, Grossi V, Landini A, Alessio MG, et al. Positive tissue transglutaminase antibodies with negative endomysial antibodies: Unresolved issues in diagnosing celiac disease. J Immunol Methods. 2021;489:112910. https://doi.org/10.1016/j.jim.2020.112910
37. Maimaris S, Schiepatti A, Conforme Torres DI, Muscia R, Gregorio V, Delogu C, et al. Biopsy□Sparing Diagnosis of Coeliac Disease Based on Endomysial Antibody Testing and Clinical Risk Assessment. Aliment Pharmacol Ther. 2025;61:1794–1804. https://doi.org/10.1111/apt.70129
38. Shoenfeld Y, Meroni PL. The general practice guide to autoimmune diseases. European Autoimmunity Standardization Initiative - EASi. Pabst Science Publishers; 2012. Lengerich, Germany. p. 1-344. ISBN: 9783899677706
39. Mäki M. The humoral immune system in coeliac disease. Baillières Clin Gastroenterol. 1995;9:231–249. https://doi.org/10.1016/0950-3528(95)90030-6
40. Hällström O. Comparison of IgA-class reticulin and endomysium antibodies in coeliac disease and dermatitis herpetiformis. Gut. 1989;30:1225–1232. https://doi.org/10.1136/gut.30.9.1225
41. Korponay-Szabó IR, Laurila K, Szondy Z, Halttunen T, Szalai Z, Dahlbom I, et al. Missing endomysial and reticulin binding of coeliac antibodies in transglutaminase 2 knockout tissues. Gut. 2003;52:199–204. https://doi.org/10.1136/gut.52.2.199
42. Nandiwada SL, Tebo AE. Testing for Antireticulin Antibodies in Patients with Celiac Disease Is Obsolete: a Review of Recommendations for Serologic Screening and the Literature. Clin Vaccine Immunol CVI. 2013;20:447–451. https://doi.org/10.1128/CVI.00568-12
43. Schwertz E, Kahlenberg F, Sack U, Richter T, Stern M, Conrad K, et al. Serologic Assay Based on Gliadin-Related Nonapeptides as a Highly Sensitive and Specific Diagnostic Aid in Celiac Disease. Clin Chem. 2004;50:2370–2375. https://doi.org/10.1373/clinchem.2004.036111
44. Fritzler MJ, Choi MY, Satoh M, Mahler M. Autoantibody Discovery, Assay Development and Adoption: Death Valley, the Sea of Survival and Beyond. Front Immunol. 2021;12:679613. https://doi.org/10.3389/fimmu.2021.679613

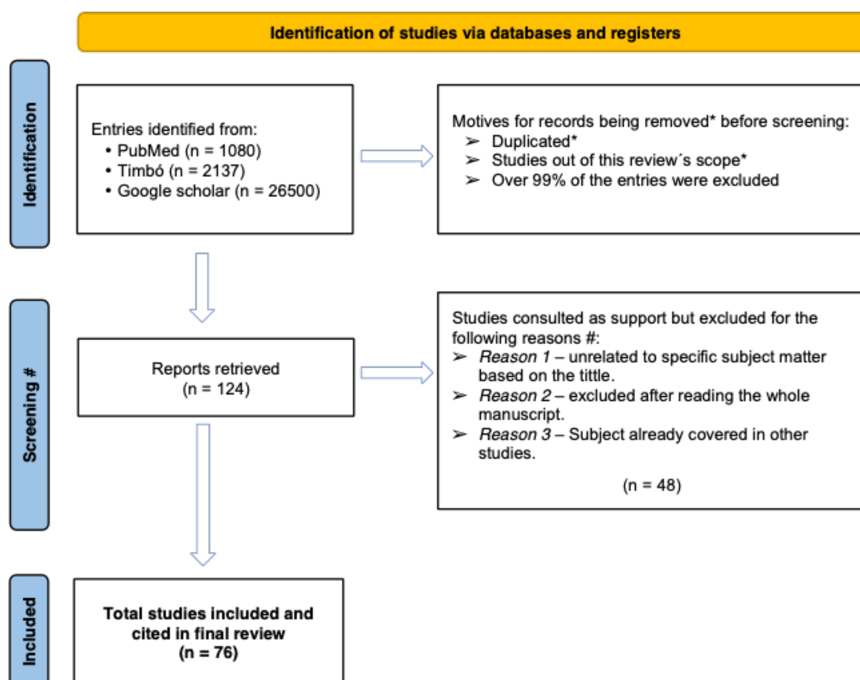
by Chemiluminescence: A Systematic Review and Meta-Analysis. Nutrients. MDPI AG; 2024;16:2427. https://doi.org/10.3390/nu16152427

45. Lewis NR, Scott BB. Systematic review: the use of serology to exclude or diagnose coeliac disease (a comparison of the endomysial and tissue transglutaminase antibody tests). Aliment Pharmacol Ther. 2006;24:47–54. https://doi.org/10.1111/j.1365-2036.2006.02967.x
46. Husby S, Koletzko S, Korponay□Szabó I, Kurppa K, Mearin ML, Ribes□Koninckx C, et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. J Pediatr Gastroenterol Nutr. 2020;70:141–156. https://doi.org/10.1097/MPG.0000000000002497
47. Rubio-Tapia A, Hill ID, Semrad C, Kelly CP, Greer KB, Limketkai BN, et al. American College of Gastroenterology Guidelines Update: Diagnosis and Management of Celiac Disease. Am J Gastroenterol. 2023;118:59–76. https://doi.org/10.14309/ajg.0000000000002075
48. Abd-Almajeed SA, Mosa ZS-A. Relative Circulating Autoantibodies IgA/IgG to Tissue Transglutaminase and Deamidated Gliadin with Celiac Disease Patients. Med Sci J Adv Res. 2025;6:1-9. https://doi.org/10.46966/msjar.v6i3.289
49. Pjetraj D, Pulvirenti A, Moretti M, Gatti S, Catassi G, Catassi C, et al. Diagnostic Accuracy of IgA Anti-Transglutaminase Assessed
50. Previtali G, Licini L, D'Antiga L, Marseglia A, Ravasio R, Nembrini F, et al. Celiac Disease Diagnosis Without Biopsy: Is a 10× ULN Antitransglutaminase Result Suitable for a Chemiluminescence Method? J Pediatr Gastroenterol Nutr. 2018;66:645–650.https://doi.org/10.1097/MPG.0000000000001773
51. Gholamrezaei R, Vaezi A, Maghool F, Jamali N, Emami MH. Does Serum Anti-tissue Transglutaminase Level Predict Duodenal Histopathology and Clinical Manifestations of Celiac Disease in Adults? Middle East J Dig Dis. 2025;17:183–189. https://doi.org/10.34172/mejdd.2025.429
52. Du N, Chang D, Wong M, Thiagarajah JR, Janssen E, Henderson LA, et al. Autoantibody signatures in children with celiac disease, juvenile idiopathic arthritis, and polyautoimmunity. JPGN Rep. 2025;jpr3.70119. https://doi.org/10.1002/jpr3.70119
53. Guz□Mark A, Kori M, Topf□Olivestone C, Weinberger R, Morgenstern S, Ziv□Sokolovskaya N, et al. Real□life Performance of Multiplex Celiac Antibody Test in the Diagnosis of Pediatric Celiac Disease. J Pediatr Gastroenterol Nutr. 2022;74:490–494. https://doi.org/10.1097/MPG.0000000000003378
54. Gomez G, Griffiths P, Fischer C, El-Shanawany T. Abstract B-079_Analytical comparison of a novel, fully automated multiplexed microarray immunoassay and routine methods for the detection of celiac disease autoantibodies in a reference laboratory in Wales. Clin Chem. 2025;71:hvaf086.477. https://doi.org/10.1093/clinchem/hvaf086.477
55. Raivio T, Korponay-Szabó I, Collin P, Laurila K, Huhtala H, Kaartinen T, et al. Performance of a new rapid whole blood coeliac test in adult patients with low prevalence of endomysial antibodies. Dig Liver Dis. 2007;39:1057–1063. https://doi.org/10.1016/j.dld.2007.09.005
56. Mendia I, Segura V, Ruiz-Carnicer Á, Coto L, Negrete M, Long JCD, et al. Rapid Anti-tTG-IgA Screening Test for Early Diagnosis of Celiac Disease in Pediatric Populations. Nutrients. 2023;15:4926. https://doi.org/10.3390/nu15234926
57. Baldas V, Tommasini A, Trevisiol C, Berti I, Fasano A, Sblattero D, et al. Development of a novel rapid non-invasive screening test for coeliac disease. Gut. 2000;47:628–631. https://doi.org/10.1136/gut.47.5.628
58. Nemec G, Ventura A, Stefano M, Leo GD, Baldas V, Tommasini A, et al. Looking for Celiac Disease: Diagnostic Accuracy of Two Rapid Commercial Assays. Am J Gastroenterol. 2006;101:1597–1600. https://doi.org/10.1111/j.1572-0241.2006.00597.x
59. Raivio T, Kaukinen K, Nemes É, Laurila K, Collin P, Kovács JB, et al. Self transglutaminase□based rapid coeliac disease antibody detection by a lateral flow method. Aliment Pharmacol Ther. 2006;24:147–154. https://doi.org/10.1111/j.1365-2036.2006.02957.x
60. Korponay-Szabó IR, Szabados K, Pusztai J, Uhrin K, Ludmány É, Nemes É, et al. Population screening for coeliac disease in primary care by district nurses using a rapid antibody test: diagnostic accuracy and feasibility study. BMJ. 2007;335:1244–1247. https://doi.org/10.1136/bmj.39405.472975.80
61. Alarida K, Harown J, Ahmaida A, Marinelli L, Venturini C, Kodermaz G, et al. Coeliac disease in Libyan children: A screening study based on the rapid determination ofanti-transglutaminase antibodies. Dig Liver Dis. 2011;43:688–691. https://doi.org/10.1016/j.dld.2011.01.002
62. Volta U, Bai JC, De Giorgio R. The role of serology in the diagnosis of coeliac disease. Gastroenterol Hepatol Bed Bench. 2023;16:118–128. https://doi.org/10.22037/ghfbb.v16i2.2713
63. Labrada OB, Reyes CV, Sánchez HNS, Otero JAS. Celiac disease in adulto. A challenge in the new millennium. Multimed Rev Médica Granma (ISSN 1028-4818). 2020;24:949–968.
64. Pumar M, Choo S, Rosenbaum J, Alex G, Ho SSC. No□Biopsy Diagnosis of Coeliac Disease in Children Without Anti□Endomysial IgA Antibody Testing: Combining Anti□Tissue Transglutaminase IgA and Anti□Deamidated Gliadin IgG Antibodies. J Paediatr Child Health. 2025;61:628–634. https://doi.org/10.1111/jpc.16801
65. Mooney PD, Kurien M, Evans KE, Chalkiadakis I, Hale MF, Kannan MZ, et al. Point-of-care testing for celiac disease has a low sensitivity in endoscopy. Gastrointest Endosc. 2014;80:456–462. https://doi.org/10.1016/j.gie.2014.02.009
66. Mäki M, Collin P. Coeliac disease. The Lancet. 1997;349:1755–1759. https://doi.org/10.1016/S0140-6736(96)70237-4
67. Al-Toma A, Volta U, Auricchio R, Castillejo G, Sanders DS, Cellier C, et al. European Society for the Study of

Coeliac Disease (ESsCD) guideline for coeliac disease and other gluten‐related disorders. United Eur Gastroenterol J. 2019;7:583–613. https://doi.org/10.1177/2050640619844125
68. Zingone F, Norman GL, Smecuol E, Maniero D, Carroccio A, Biagi F, et al. Utilizing both IgA tissue transglutaminase and IgG-deamidated gliadin peptide antibodies offers accurate celiac disease diagnosis without duodenal biopsy. Dig Liver Dis. 2025;57:609–615. https://doi.org/10.1016/j.dld.2024.10.010
69. Gambino CM, Agnello L, Del Ben F, Ciaccio AM, Milano S, Vassallo R, et al. Evaluating the performance of a multiparametric IgA assay for celiac disease diagnosis. Clin Chem Lab Med CCLM. 2025;63:1-7. https://doi.org/10.1515/cclm-2025-0705
70. Parrinello G, Da Re M, Grizzo F, Camelliti S, Cozzi M, Marinoni F, et al. Diagnostic accuracy of a novel point‐of‐care test for simultaneous detection of anti‐transglutaminase IgA and anti‐deamidated gliadin IgG antibodies. J Clin Lab Anal. 2024;38:e25003. https://doi.org/10.1002/jcla.25003
71. Mpakosi A, Kaliouli-Antonopoulou C, Cholevas V, Cholevas S, Tzouvelekis I, Mironidou-Tzouveleki M, et al. Challenges in the Pediatric Celiac Disease Diagnosis: An Up-to-Date Review. Diagnostics. 2025;15:2392. https://doi.org/10.3390/diagnostics15182392
72. Wolf J, Petroff D, Richter T, Auth MKH, Uhlig HH, Laass MW, et al. Validation of Antibody-Based Strategies for Diagnosis of Pediatric Celiac Disease Without Biopsy. Gastroenterology. 2017;153:410-419.e17. https://doi.org/10.1053/j.gastro.2017.04.023
73. Chokkalla AK, Parham MM, Fishman DS, Devaraj S. Path Towards Biopsy-Free Diagnosis of Celiac Disease in Pediatric Patients. Clin Chim Acta. 2024;557:117891. https://doi.org/10.1016/j.cca.2024.117891
74. Pavelescu LA, Sabau ID, Sanda-Dira G, Iacata AA, Curici A. Serological, Genetic, and Biochemical Insights into Celiac Disease Diagnosis and Vitamin D Deficiency in Romanian Children: A Comprehensive Cohort Study. Int J Mol Sci. 2025;26:6251. https://doi.org/10.3390/ijms26136251
75. Dreyfuss M, Getz B, Lebwohl B, Ramni O, Underberger D, Ber TI, et al. A machine learning tool for early identification of celiac disease autoimmunity. Scientific Reports. 2024;14:30760. https://doi.org/10.1038/s41598-024-80817-0
76. Piroozkhah M, Kahangi MF, Asri N, Piroozkhah M, Moradi N, Nazemalhosseini-Mojarad E, et al. Toward Precision Medicine in Celiac Disease: Emerging Roles of Multiomics and Artificial Intelligence. Tuccari G, editor. Anal Cell Pathol. 2025;2025:6570955. https://doi.org/10.1155/ancp/6570955

Supplementary file

Supplementary Figure 1: Flowchart according to the PRISMA-S method for the selection, screening, and inclusion of studies in the review.



Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review Article

Interrelationship between Serum FGF21, GDF15, and Microalbuminuria as Predictive Biomarkers for Early Detection of Diabetic Retinopathy in Type 2 Diabetes Mellitus

Mustakin, M.Kes, Sp.PK[1*]

[1]Department of Clinical Pathology, Faculty of Medicine, Hasanuddin University, Makassar, Indonesia

Article Info

****Corresponding Author:***

Mustakin, M.Kes, Sp.PK
Department of Clinical Pathology, Faculty of Medicine
Hasanuddin University, Makassar, Indonesia
E-mail: mustakimduharing@ymail.com

Keywords

FGF21, GDF15, Microalbuminuria, Diabetic retinopathy, Biomarkers, Mitochondrial stress, Endothelial dysfunction, Type 2 diabetes mellitus

Abstract

Diabetic retinopathy (DR) is a leading cause of preventable visual loss in type 2 diabetes mellitus (T2DM), yet current screening largely detects disease after microvascular damage has occurred. This review synthesizes emerging evidence that three measurable biomarkers - fibroblast growth factor-21 (FGF21), growth differentiation factor-15 (GDF15), and microalbuminuria -capture complementary dimensions of early DR pathobiology. FGF21 reflects hepatometabolic stress and adaptive signaling via the β-Klotho–FGFR1 axis; GDF15 mirrors mitochondrial/integrated-stress-response activation with systemic effects through the glial cell-derived neurotrophic factor family receptor-α-like (GFRAL) REarranged during Transfection (RET); and microalbuminuria indicates systemic endothelial dysfunction and increased microvascular permeability. We narratively integrate cross-sectional, cohort, and meta-analytic studies in human T2DM populations (excluding type 1 diabetes and animal experiments) and map these signals onto a unified metabolic–renal–retinal framework. Across studies, higher circulating FGF21 and GDF15 associate with DR presence and severity, while microalbuminuria correlates with DR grade and predicts progression independent of glycemic control. We propose a tri-biomarker model in which chronic lipotoxicity and oxidative stress elevate FGF21/GDF15, drive endothelial injury detectable as microalbuminuria, and together forecast early retinal microangiopathy. Translational implications include risk stratification before funduscopic changes, therapy monitoring for metabolic/mitochondrial targets, and development of multimarker algorithms alongside OCTA imaging. To our knowledge, this is the first review to integrate FGF21, GDF15, and microalbuminuria as a coherent predictive axis for early DR in T2DM, highlighting priorities for longitudinal validation and population-specific cut-offs.

Background and Rationale

Diabetic retinopathy (DR) is a leading cause of preventable blindness among patients with type 2 diabetes mellitus (T2DM). Despite advances in retinal screening technologies, most cases of DR are still detected only after clinically apparent microvascular damage has occurred [1,2].

DR is not solely the result of chronic hyperglycemia; rather, it involves a complex pathophysiological process characterized by metabolic stress, mitochondrial dysfunction, neurovascular unit impairment, inflammation, and systemic endothelial injury [3–5,6].

Among circulating biomarkers, fibroblast growth factor-21 (FGF21) and growth differentiation factor-15 (GDF15) have emerged as key indicators of metabolic and mitochondrial stress in T2DM. Elevated FGF21 levels have been associated with the presence and severity of DR [6–10], whereas increased GDF15 concentrations have been consistently observed in patients with T2DM exhibiting microvascular complications, including DR [11–14]. Both biomarkers rise in response to lipotoxicity, oxidative stress, and activation of the integrated stress response (ISR) pathway.

In parallel, microalbuminuria - reflected by an increased urinary albumin-to-creatinine ratio (UACR) - serves as an established marker of systemic endothelial injury and enhanced microvascular permeability. Numerous studies have demonstrated strong associations between albuminuria and both the presence and progression of DR [15–18].

Molecular Basis of Fibroblast Growth Factor-21 (FGF21) in Retinal Microangiopathy

Fibroblast growth factor-21 (FGF21) is a hepatokine that increases in response to fasting, lipotoxic stress, and chronic hyperglycemia, functioning to regulate glucose and lipid metabolism through the β-Klotho–fibroblast growth factor receptor-1 (FGFR1) complex. Under conditions of chronic metabolic stress, elevated reactive oxygen species stimulate FGF21 expression as a compensatory mechanism to maintain mitochondrial homeostasis [6–9].

Consistent elevations in circulating FGF21 have been reported in patients with T2DM across various stages of diabetic retinopathy [6–10]. Mechanistically, FGF21 exhibits a dual role. In the early phase, it acts as a metabolic protector by enhancing insulin sensitivity, increasing adiponectin secretion, and reducing the formation of reactive oxygen species. However, during prolonged metabolic overload, FGF21 resistance develops - characterized by impaired β-Klotho–FGFR1 signaling and paradoxically elevated FGF21 concentrations [9]. This maladaptive state parallels retinal hypoxia, endothelial dysfunction, and oxidative imbalance, which represent hallmark features of early retinal microangiopathy.

Recent meta-analyses strengthen this association. Basir et al. reported that higher circulating FGF21 levels were significantly linked to the risk and severity of diabetic retinopathy, a finding further confirmed by another meta-analysis demonstrating consistent FGF21 elevation across multiple disease stages [11,12]. Biologically, FGF21 is involved in retinal lipid regulation via activation of peroxisome proliferator-activatedreceptor-α (PPAR-α), modulation of oxidative pathways, and suppression of apoptosis related to endoplasmic reticulum stress. Overall, FGF21 reflects both compensatory and maladaptive processes in systemic and retinal microvascular metabolism, underscoring its potential as an early biomarker to identify individuals at increased risk for diabetic retinopathy.

Growth Differentiation Factor-15 (GDF15) and Mitochondrial Stress Signaling in Retinopathy

Growth differentiation factor-15 (GDF15) is a member of the transforming growth factor-β (TGF-β) superfamily and functions as a mitochondrial distress signal. Its expression increases through the integrated stress response (ISR) pathway, primarily via activation of CCAAT/enhancer-binding protein homologous protein (CHOP), activating transcription factor-4 (ATF4), and p53 during oxidative stress or endoplasmic reticulum stress [13–15]. Several clinical studies have consistently reported elevated serum GDF15 levels in patients with T2DM who have diabetic retinopathy [13–15].

Unlike fibroblast growth factor-21 (FGF21), which acts predominantly in peripheral tissues, GDF15 exerts systemic effects through the glial cell-derived neurotrophic factor family receptor-α-like (GFRAL) and REarranged during Transfection (RET) receptor complex located in the brainstem. Activation of the GFRAL–RET pathway modulates inflammation, energy metabolism, and systemic stress responses. High circulating GDF15 levels in patients with T2DM reflect mitochondrial stress and chronic endothelial injury, both of which contribute to retinal microvascular degeneration [16].

At the molecular level, GDF15 helps suppress excessive production of reactive oxygen species (ROS); however, its sustained elevation indicates failure of cellular adaptive mechanisms in the face of prolonged metabolic stress. This positions GDF15 within an adaptive–maladaptive continuum, whereby simultaneous increases in GDF15 and FGF21 occur as the system transitions from metabolic compensation toward progressive vascular injury.

Recent evidence also highlights the involvement of GDF15 in renal microangiopathy, with elevated GDF15 levels observed in T2DM patients exhibiting microalbuminuria. This finding aligns with data showing higher GDF15 concentrations in patients who present with both diabetic retinopathy and impaired renal function [17,18]. Collectively, this dual expression pattern across retinal and renal tissues reinforces the role of GDF15 as a systemic biomarker reflecting microvascular dysfunction in type 2 diabetes.

Microalbuminuria as a Systemic Marker of Endothelial Injury

Microalbuminuria, defined as urinary albumin excretion of 30–300 mg/day, is one of the earliest markers of microvascular dysfunction in T2DM [19]. This condition reflects increased permeability of the glomerular capillary wall and indicates the presence of systemic endothelial injury. Numerous cross-sectional and cohort studies have demonstrated that even mild elevations in albuminuria are significantly associated with

both the presence and severity of diabetic retinopathy (DR) [19–21].

Pathophysiologically, chronic hyperglycemia, the accumulation of advanced glycation end-products (AGEs), and oxidative stress lead to disruption of the endothelial glycocalyx and neurovascular unit dysfunction, resulting in vascular injury that affects both the renal glomerulus and the retinal microcirculation [22]. Large population-based studies further strengthen this association. The Singapore Malay Eye Study reported that microalbuminuria independently increases the risk of any-stage DR [22], whereas the Sankara Nethralaya Diabetic Retinopathy Epidemiology and Molecular Genetics Study (SN-DREAMS) demonstrated that albuminuria predicts proliferative DR even after adjustment for glycemic control and blood pressure [21]. Therefore, albuminuria is not merely a renal manifestation butfunctions as a systemic vascular biomarker integrating metabolic, inflammatory, and hemodynamic stress.

In recent years, the relationship among fibroblast growth factor-21 (FGF21), growth differentiation factor-15 (GDF15), microalbuminuria, and diabetic retinopathy has been increasingly explored across diverse study designs. Consistently, elevated FGF21 and GDF15 reflect metabolic stress and mitochondrial dysfunction, whereas microalbuminuria indicates systemic endothelial injury. Although inter-study variations exist in methodology and population characteristics, the majority of findings report positive associations between these biomarkers and the presence or progression of DR. Collectively, this evidence supports their roles as early indicators of metabolic–microvascular dysregulation preceding clinically detectable retinal damage.

Table 1: summarizes the key clinical and meta-analytic studies evaluating the associations between serum FGF21, GDF15, and microalbuminuria with the presence and severity of diabetic retinopathy in patients with type 2 diabetes mellitus.

Author (Year)	Sample size / Study design	Biomarker(s) assessed	Main findings	Ref
Jin S et al. (2021)	312 T2DM patients / cross-sectional	Serum FGF21	Elevated FGF21 significantly associated with sight-threatening diabetic retinopathy.	[1]
Heidari Z, Hasanpour M (2021)	210 T2DM patients / case–control	Serum FGF21	FGF21 positively correlated with DR severity, independent of HbA1c and lipid profile.	[2]
Jung CH et al. (2017)	1,110 T2DM subjects / cohort	Serum FGF21	U-shaped association between FGF21 levels and microvascular complications, including DR.	[3]
Esteghamati A et al. (2016)	140 T2DM patients / case–control	Serum FGF21	Circulating FGF21 levels significantly higher in DR compared with non-DR patients.	[4]
Lee CH et al. (2023)	438 T2DM patients / prospective	FGF21, AFABP, PEDF	Elevated FGF21 predicts progression to sight-threatening diabetic retinopathy.	[5]
Basir H et al. (2024)	Meta-analysis (10 studies; n≈3,400)	Serum FGF21	Meta-analysis demonstrates a strong association between elevated FGF21 levels and DR risk.	[6]
Jiang Y et al. (2024)	Meta-analysis (9 studies)	FGF21, Klotho	Higher FGF21 and lower Klotho levels associated with increased DR severity.	[7]
Niu Y et al. (2021)	402 T2DM patients / cross-sectional	Serum GDF15	GDF15 significantly elevated in patients with DR and correlates with disease severity.	[11]

Author (Year)	Sample size / Study design	Biomarker(s) assessed	Main findings	Ref
Chung JO et al. (2020)	312 T2DM patients / cross-sectional	Plasma GDF15	Higher plasma GDF15 levels associated with presence and severity of DR.	[12]
Billeson K et al. (2024)	214 diabetic subjects / case–control	Serum GDF15, MMP-3	GDF15 levels associated with diabetic microvascular complications, including retinopathy.	[13]
Feng N et al. (2026)	3369 adults diabetes / cross-sectional	UACR	Higher urinary albumin-to-creatinine ratio positively associated with DR and may serve an early predictive DR	[18]
Paterson N et al. (2021)	1,883 subjects T2DM / cross-sectional case control(UK Biobank)	Microalbuminuria	Albuminuria is associated with alterations in retinal microvascular parameters.	[19]
Dash S et al. (2022)	250 T2DM patients / cross-sectional	Albuminuria	Severity of albuminuria positively correlates with the grade of diabetic retinopathy.	[20]
Rani PK et al. (2011)	1,411 T2DM patients / cohort (SN-DREAMS)	Albuminuria	Albuminuria predicts proliferative DR independent of HbA1c and blood pressure.	[21]
Shahrir NF et al. (2022)	343 T2DM patients / cross-sectional study	Microalbuminuria	Microalbuminuria independently associated with diabetic retinopathy in Malay population.	[22]

Integrative Mechanism

The Metabolic–Renal–Retinal Axis

The interplay among fibroblast growth factor-21 (FGF21), growth differentiation factor-15 (GDF15), and microalbuminuria represents a unified pathophysiological pathway that bridges hepatic metabolism, mitochondrial stress, and endothelial integrity in *type 2 diabetes mellitus (T2DM)* [6,13,19]. This biological framework is conceptualized as the Metabolic–Renal–Retinal Axis (Figure 1).

Metabolic Overload

Chronic lipotoxicity and excessive fatty acid flux activate peroxisome proliferator-activated receptor-α (PPAR-α) and mitochondrial stress pathways, leading to the upregulation of FGF21 and GDF15 as adaptive metabolic hormones [6,12].

Mitochondrial Stress and the Integrated Stress Response (ISR)

Impaired oxidative phosphorylation and reactive oxygen species accumulation trigger the integrated stress response, involving CHOP, ATF4, and p53 activation, which in turn induces systemic elevation of GDF15 as a mitochondrial distress signal [13–15].

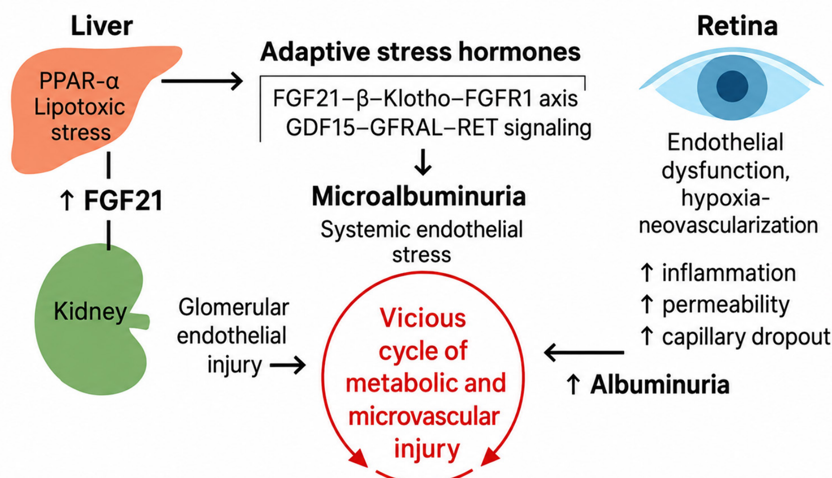
Endothelial Dysfunction

Oxidative and inflammatory insults disrupt the endothelial glycocalyx and tight junctions, causing urinary albumin leakage and increased retinal microvascular fragility. This shared mechanism underlies the parallel emergence of microalbuminuria and early diabetic retinopathy lesions [19–22].

Feed-Forward Loop

Elevated FGF21 and GDF15 further modulate inflammation, vascular permeability, and metabolic stress responses, creating a self-reinforcing cycle that accelerates microvascular injury in both renal and retinal tissues [11,16]. Collectively, this integrative framework illustrates how metabolic overload stimulates hepatic FGF21 secretion, mitochondrial stress drives GDF15 upregulation, and endothelial damage manifests as microalbuminuria. Together, these interlinked pathways form a feed-forward loop that amplifies systemic inflammation, oxidative stress, and microvascular permeability. This mechanistic model provides the biological foundation for a multi-biomarker strategy aimed at early prediction and detection of diabetic retinopathy in individuals with T2DM.

Figure 1: The Metabolic–Renal–Retinal Axis.



Source: Created by the authors using BioRender (www.biorender.com), adapted and synthesized from previously published studies [11,12,18,22].

This Figure illustrates the integrated roles of fibroblast growth factor-21 (FGF21) and growth differentiation factor-15 (GDF15) as circulating metabolic stress hormones, and microalbuminuria as a systemic marker of endothelial dysfunction. Chronic lipotoxicity and oxidative stress promote hepatic FGF21 secretion and systemic GDF15 elevation, which converge to impair endothelial integrity within renal and retinal microcirculation. These mechanisms generate a feed-forward loop that amplifies metabolic stress, vascular permeability, and microvascular injury in type 2 diabetes mellitus.

Clinical and Translational Implications

The combined assessment of fibroblast growth factor-21 (FGF21), growth differentiation factor-15 (GDF15), and microalbuminuria offers a promising approach for enhancing the early detection of diabetic retinopathy (DR). As a non-invasive, low-cost biomarker triad, these markers capture complementary aspects of metabolic stress, mitochondrial dysfunction, and endothelial injury, providing a more sensitive reflection of microvascular risk than glycemic indices alone [11,13,19]. Diabetic retinopathy remains one of the leading causes of vision impairment worldwide and represents a complex microvascular complication driven by chronic metabolic dysregulation in diabetes mellitus [24].

From a clinical perspective, this integrative biomarker framework may provide several practical applications for improving risk assessment and management in patients with type 2 diabetes mellitus. Early identification of individuals at risk of developing diabetic retinopathy is essential for preventing irreversible retinal damage. Biomarker-based profiling using FGF21, GDF15, and microalbuminuria may help identify patients experiencing early systemic microvascular stress before the appearance of overt retinal lesions. This approach may complement established international classification systems used to stage diabetic retinopathy severity and guide clinical monitoring strategies [25]. Recent advances in screening programs emphasize the need for combining systemic biomarkers with retinal imaging technologies to enhance early detection of retinal microvascular damage. Novel screening frameworks incorporate digital retinal photography and emerging imaging modalities such as optical coherence tomography angiography (OCTA) to improve diagnostic sensitivity and population-based screening efficiency [26]. Understanding the biological mechanisms underlying diabetic retinopathy is critical for improving prevention strategies. Chronic hyperglycemia induces metabolic stress, oxidative damage, and endothelial dysfunction that contribute to progressive retinal microvascular injury. These mechanisms have long been recognized as key contributors to the development and progression of diabetic retinopathy [27]. Circulating levels of FGF21 and GDF15 may provide insights into systemic metabolic stress and mitochondrial dysfunction during treatment interventions targeting insulin resistance or lipid metabolism. Experimental studies suggest that GDF15 may also play regulatory roles in inflammation, cellular stress signaling, and tissue remodeling processes associated with metabolic disease [28]. Microalbuminuria represents a widely recognized marker of endothelial dysfunction and systemic microvascular injury in diabetes. Several studies have demonstrated that increased urinary albumin excretion correlates with the presence and severity of diabetic microvascular complications, including diabetic retinopathy [29]. Retinal vascular endothelial cell damage and neuroretinal degeneration are increasingly recognized as central mechanisms contributing to diabetic retinopathy progression. Endothelial dysfunction disrupts retinal microvascular integrity and promotes inflammatory and neurodegenerative processes within the retina [30]. Microalbuminuria has also been proposed as an early clinical indicator of systemic vascular injury in patients with type 2 diabetes mellitus and may precede the development of overt diabetic complications. Emerging therapeutics targeting thesepathways, including FGF21 analogs and GDF15–GFRAL axis modulators, demonstrate potential in reducing systemic metabolic stress and inflammation. However, their application in DR remains exploratory, and longitudinal validation is required to determine whether modifying these biomarkers translates into reduced microvascular injury [12,16,22].

Taken together, these translational insights highlight the potential of multi-biomarker models to complement current screening tools and advance personalized approaches to the prevention of diabetic retinopathy.

Future Prospects and Conclusion

Microalbuminuria, fibroblast growth factor-21 (FGF21), and growth differentiation factor-15 (GDF15) collectively reflect the metabolic, mitochondrial, and endothelial stress underpinning microvascular injury in type 2 diabetes mellitus (T2DM). Their combined assessment may substantially improve early diagnosis and prediction of diabetic retinopathy (DR), especially during subclinical stages when conventional ophthalmic screening has limited sensitivity [11,13,19].

Future research should prioritize large-scale longitudinal studies to establish standardized cut-off values for each biomarker, determine temporal trajectories in relation to DR progression, and examine gene - environment interactions influencing biomarker expression. Population-based studies across diverse ethnic groups - including Indonesia's Bugis-Makassar population - are particularly important to account for genetic variation, metabolic phenotypes, and differing susceptibilities to microvascular complications.

To the best of current knowledge, the present review is the first to integrate FGF21, GDF15, and microalbuminuria into a unified Metabolic–Renal–Retinal Axis, offering a comprehensive biological model that links metabolic overload, mitochondrial stress, and endothelial dysfunction to early retinal microangiopathy. This tri-biomarker framework provides mechanistic insights that extend beyond glycemic control and may complement existing screening strategies by identifying high-risk individuals earlier.

Further translational research is required to validate this integrative model, define population-specific thresholds, and evaluate whether modifying these biomarkers - through metabolic, mitochondrial, or vascular-targeted therapies - can reduce microvascular injury and ultimately prevent vision-threatening diabetic retinopathy.

Acknowledgments

The authors express sincere appreciation to the Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Universitas Hasanuddin, for academic and technical support.

Funding Statement

No external funding was received for this review.

Conflict of Interest

The authors declare no conflict of interest related to this manuscript.

Ethics Statement

No ethical approval was required for this review article as it did not involve direct patient participation or identifiable human data.

Artificial Intelligence Assistance Declaration

Artificial intelligence assistance was utilized in the preparation of this manuscript exclusively for non-creative technical tasks, including language polishing, structural refinement, and reference formatting. The AI model used was OpenAI's ChatGPT(GPT-5 model), accessed through a subscription-based interface. No content generation, data interpretation, or scientific conclusion was made autonomously by the AI. All critical analyses, conceptual frameworks, and interpretations were independently developed, reviewed, and validated by the authors. The use of AI tools complied with the journal's ethical standards and current COPE guidelines on responsible use of generative language models in scientific writing.

References

1. Jin S, Xia N, Han L. Association between serum fibroblast growth factor 21 level and sight-threatening diabetic retinopathy in Chinese patients with type 2 diabetes. BMJ Open Diabetes Research & Care. 2021;9(1):e002126. https://doi.org/10.1136/bmjdrc-2021-002126
2. Heidari Z, Hasanpour M. The serum fibroblast growth factor 21 is correlated with retinopathy in patients with type 2 diabetes mellitus. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2021;15(6):102296. https://doi.org/10.1016/j.dsx.2021.102296
3. Jung CH, Jung SH, Kim BY, Kim CH, Kang SK, Mok JO. The U-shaped relationship between fibroblast growth factor 21 and microvascular complication in type 2 diabetes mellitus. Journal of Diabetes and its Complications. 2017;31(1):134. https://doi.org/10.1016/j.jdiacomp.2016.10.017
4. Esteghamati A, Momeni A, Abdollahi A, Khandan A, Afarideh M, Noshad S, et al. Dosage sérique du facteur de croissance des fibroblastes 21 chez les patients atteints de rétinopathie diabétique de type 2. Annales d'Endocrinologie. 2016;77(5):586. https://doi.org/10.1016/j.ando.2016.01.005
5. Lee CH, Lui DTW, Cheung CY, Fong CHY, Yuen MMA, Woo YC, et al. Circulating AFABP, FGF21, and PEDF levels as prognostic biomarkers of sight-threatening diabetic retinopathy. Journal of Clinical Endocrinology and Metabolism. 2023;108(9):E799. https://doi.org/10.1210/CLINEM/DGAD112
6. Basir H, Nugrahani ASD, Aman AM, Bakri S, Rasyid H, Umar H, et al. The association between fibroblast growth factor 21 with diabetic retinopathy among type 2 diabetes mellitus patients: a systematic review, meta-analysis, and meta-regression. PeerJ. 2024;12(12). https://doi.org/10.7717/PEERJ.18308
7. Jiang Y, Zhang W, Xu Y, Zeng X, Sun X. Relationship of fibroblast growth factor 21, Klotho, and diabetic retinopathy: a meta-analysis. Frontiers in Endocrinology. 2024;15. https://doi.org/10.3389/FENDO.2024.1390035
8. Wei L, Sun X, Fan C, Li R, Zhou S, Yu H. The pathophysiological mechanisms underlying diabetic retinopathy. Frontiers in Cell and Developmental Biology. 2022;10:963615. https://doi.org/10.3389/fcell.2022.963615

9. Shakthiya T, Chand L, Annamalai R. A recent update on candidate biomarkers in the pathogenesis of diabetic retinopathy. The Open Biomarkers Journal. 2025;15(1). https://doi.org/10.2174/0118753183372247250213110930
10. Niu Y, Ding X, Yang Z, Han Y, Zhang H, Zhang X, et al. The relationship between circulating growth differentiation factor-15 levels and diabetic retinopathy in patients with type 2 diabetes. Frontiers in Endocrinology. 2021;12:627395. https://doi.org/10.3389/fendo.2021.627395
11. Chung JO, Park SY, Cho DH, Chung DJ, Chung MY.
12. Relationship between plasma growth differentiation factor-15 levels and diabetic retinopathy in individuals with type 2 diabetes.Scientific Reports. 2020;10:20568. https://doi.org/10.1038/s41598-020-77584-z
13. Billeson K, Baldimtsi E, Wahlberg J, Whiss PA. Growth Differentiation Factor 15 and Matrix Metalloproteinase 3 in Plasma as Biomarkers for Neuropathy and Nephropathy in Type 1 Diabetes. Int J Mol Sci. 2024;25(13):7328. https://doi.org/10.3390/ijms25137328
14. Berezin AE. Diabetes mellitus related biomarker: The predictive role of growth-differentiation factor-15. Diabetes & Metabolic Syndrome: Clinical Research & Reviews. 2016;10(1):S154. https://doi.org/10.1016/j.dsx.2015.09.016
15. Torm MEW, Dorweiler TF, Fickweiler W, et al. Frontiers in diabetic retinal disease. Journal of Diabetes and its Complications. 2023;37(2):108386. https://doi.org/10.1016/j.jdiacomp.2022.108386
16. Soelistijo SA, Ardiany D, Adiwinoto RD, Pranoto A, Widjaja SA.Albuminuria as a risk factor for diabetic retinopathy in Indonesian type 2 diabetes mellitus patients.
17. Rom J Diabetes Nutr Metab Dis. 2024;31(2):170–175. https://doi.org/10.46389/rjd-2024-1446
18. Supraja P, Hemalatha K, Suneetha P.Correlation of microalbuminuria and diabetic retinopathy: a retrospective study.Int J Res Med Sci. 2024;12(11):4185–4192. https://doi.org/10.18203/2320-6012.ijrms20243370
19. Antwi-Boasiako C, Obeng KA, Amissah-Arthur KN, Musah L, Abaidoo B, Awuviri HW, et al. Association between albuminuria and retinal microvascular dysfunction in type 2 diabetes with and without hypertension. Diabetes Epidemiology and Management. 2023;11:100139. https://doi.org/10.1016/j.deman.2023.100139
20. Feng N, Wang YX, Tian XJ, Zhao SF, Du LJ, Feng BQ, et al. Positive association between urinary albumin-to-creatinine ratio and diabetic retinopathy in American adults: a cross-sectional study. Medicine (Baltimore). 2026;105:e47439. https://doi.org/10.1097/MD.0000000000047439
21. Paterson EN, Cardwell CR, MacGillivray TJ, Trucco E, Doney ASF, Foster PJ, et al. Investigation of associations between retinal microvascular parameters and albuminuria in UK Biobank: a cross-sectional case-control study. BMC Nephrol. 2021;22:72. https://doi.org/10.1186/s12882-021-02273-6
22. Dash S, Chougule A, Mohanty S. Correlation of albuminuria and diabetic retinopathy in type-II diabetes mellitus patients. Cureus. 2022;14(2):e21927. https://doi.org/10.7759/cureus.21927
23. Rani PK, Raman R, Gupta A, Pal SS, Kulothungan V, Sharma T. Albuminuria and diabetic retinopathy in type 2 diabetes mellitus: Sankara Nethralaya Diabetic Retinopathy Epidemiology and Molecular Genetic Study (SN-DREAMS, Report 12). Diabetol Metab Syndr. 2011;3:9. https://doi.org/10.1186/1758-5996-3-9
24. Shahrir NF, Aziz NRAA, Ahmad FL, Muzaid NA, Samat F, Syed Ghazaili SNA, et al. Determinants of microalbuminuria among type 2 diabetes mellitus patients in Kuala Selangor district: A cross-sectional study. Malays Fam Physician. 2022;17(3):53–63. https://doi.org/10.51866/oa.122
25. Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. The Lancet. 2010;376(9735):124. https://doi.org/10.1016/S0140-6736(09)62124-3
26. Wilkinson CP, Ferris FL, Klein RE, Lee PP, Agardh CD, Davis M, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology. 2003;110(9):1677. https://doi.org/10.1016/S0161-6420(03)00475-5
27. Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, et al. Screening for diabetic retinopathy: new perspectives and challenges. The Lancet Diabetes and Endocrinology. 2020;8(4):337. https://doi.org/10.1016/S2213-8587(19)30411-5
28. Aiello LM. Perspectives on diabetic retinopathy. American Journal of Ophthalmology. 2003;136(1):122. https://doi.org/10.1016/S0002-9394(03)00219-8
29. Sawant H, Borthakur A. Disease-specific novel role of growth differentiation factor 15 in organ fibrosis. International Journal of Molecular Sciences. 2025;26(12):5713. https://doi.org/10.3390/ijms26125713
30. Yadav S, Singhal S, Shashwat K, Saklani A, Mittal S, Anand H. Study of correlation of urine albumin creatinine ratio with diabetic retinopathy in type 2 diabetes mellitus patients. International Journal of Scientific Development and Research. 2024;9(11): Available from https://www.ijsdr.org/papers/IJSDR2411003.pdf
31. Mrugacz M, Bryl A, Zorena K. Retinal vascular endothelial cell dysfunction and neuroretinal degeneration in diabetic patients. Journal of Clinical Medicine. 2021;10(3):458. https://doi.org/10.3390/jcm10030458
32. Asghar S, Asghar S, Mahmood T, Bukhari SMH, Mumtaz MH, Rasheed A. Microalbuminuria as the Tip of Iceberg in Type 2 Diabetes Mellitus: Prevalence, Risk Factors, and Associated Diabetic Complications. Cureus. 2023;15(8):e43190. https://doi.org/10.7759/cureus.43190

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

Acute Intermittent Porphyria with a Secondary Porphyria Cutanea Tarda–like Biochemical Pattern in a patient with co-morbid Human Immunodeficiency Virus infection

Kagiso M. Masemola[1,2*], Kgaogelo R. Masemola[2,3], Mogomotsi Dintshi[1,2], Taryn Pillay[2,3]

[1]Department of Chemical Pathology, School of Pathology, University of the Witwatersrand, Johannesburg (South Africa)
[2]National Health Laboratory Services (South Africa)
[3]School of Laboratory Medicine, Medical Science, College of Health Sciences, University of KwaZulu-Natal (South Africa)

Article Info

Corresponding Author:
Kagiso Masemola
University of the Witwatersrand
Johannesburg, South Africa, 2193
Email: kagiso.masemola@nhls.ac.za

Keywords

Heme, Porphyria, Acute intermittent porphyria (AIP), Porphyria cutanea tarda (PCT)

Abstract

Introduction: Porphyrias are rare inherited disorders of heme synthesis caused by reduced activity of one of the eight enzymes in the pathway. When early precursors accumulate, acute hepatic porphyrias occur, which present as severe neurovisceral attacks. When later, light-sensitive porphyrins accumulate, they cause cutaneous forms with chronic photosensitivity, blistering and skin fragility. Although these patterns are usually distinct, acute, and cutaneous features may appear together. This may reflect two separate genetic defects, but more often occurs when oxidative stress, such as from iron overload, chronic infection, or drugs, secondarily inhibits the fifth enzyme, uroporphyrinogen decarboxylase. This creates a mixed porphyrin pattern in which the usual distinctions between acute and cutaneous porphyrias are blurred, making diagnosis challenging.

Case Presentation: A 22-year-old woman living with human immunodeficiency virus (HIV) infection presented with recurrent abdominal pain during the luteal phase of menstruation, autonomic instability, and limb weakness that improved following heme arginate therapy. Urine porphobilinogen was markedly elevated, and hydroxymethylbilane synthase (*HMBS)* gene sequencing confirmed acute intermittent porphyria (AIP). However, her porphyrin studies also demonstrated pronounced urinary uroporphyrin elevation together with faecal isocoproporphyrins, a pattern characteristic of porphyria cutanea tarda (PCT).

Discussion: The findings are consistent with AIP complicated by secondary hepatic inhibition of uroporphyrinogen decarboxylase. In this patient, secondary uroporphyrinogen decarboxylase (UROD) inhibition was likely driven by persistent HIV viraemia, antiretroviral-associated hepatic oxidative stress, and chronic inflammatory activation which are well-recognised risk factors for acquired PCT-like biochemical patterns.

Conclusion
This case demonstrates how co-morbidities can modify classical porphyria biochemical patterns and reinforces the need for integrated urine, plasma, and faecal interpretation, especially when biochemical profiles appear mixed. Clinically, patients living with HIV who have porphyria may benefit from closer monitoring for factors that increase liver stress, especially ongoing viraemia, hepatotoxic or porphyrinogenic drugs, alcohol use and iron overload. When virological control is poor, the risk of secondary UROD inhibition and mixed biochemical findings may increase. Regular review of HIV virological control, liver function, medication safety and porphyrin profiles may therefore help prevent recurrent attacks and reduce diagnostic confusion.

Introduction
Porphyrias are disorders caused by reduced activity of one of the eight enzymes in the heme biosynthesis pathway [1]. When an enzyme is impaired, intermediates before the metabolic block accumulate [1,2]. These accumulating precursors, rather than reduced heme production, cause the clinical features [2,3]. The pattern of precursor accumulation determines the clinical phenotype [1,3].
Acute porphyrias such as AIP, from reduced porphobilinogen deaminase activity, present with severe neurovisceral attacks [2,3]. In contrast, cutaneous porphyrias such as PCT, due to reduced UROD activity, cause blistering photosensitivity [2,4]. Some forms such as variegate porphyria show both acute and cutaneous manifestations [1].
Acute attacks are medical emergencies because rapid rises in delta-aminolevulinic acid (ALA) and porphobilinogen (PBG) disrupt neuronal function [3]. Acute porphyria presents with severe abdominal and autonomic symptoms, which may progress to seizures, behavioural changes or respiratory failure [2,3].
Cutaneous porphyrias occur when photoactive porphyrins accumulate in the skin, causing blistering and fragility on light exposure [2,4,5].
Although porphyrias are usually considered separate entities, mixed biochemical patterns do occur [1,3]. Simultaneous accumulation of early and late intermediates may indicate true dual porphyria or a single inherited defect complicated by secondary reduction of uroporphyrinogen decarboxylase activity. In people with HIV, drug-induced cytochrome P450 induction, hepatotoxicity and disordered iron metabolism can create this secondary block [5,6].
This report describes a young woman with genetically confirmed AIP and a PCT biochemical pattern, illustrating the diagnostic complexity in HIV-related hepatic stress.

Case Description
A 22-year-old Black South African woman with HIV infection, diagnosed three years prior, was receiving fixed-dose, once daily, antiretroviral therapy (ART) in the form of tenofovir (TDF) 300 mg, lamivudine (3TC) 300 mg and dolutegravir 50 mg (TLD). She presented with recurrent, cyclical episodes of acute abdominal pain, tachycardia, vomiting, and progressive limb weakness that reliably coincided with the luteal phase of hermenstrual cycle. She denied alcohol use, prolonged fasting, or recreational drugs.

Initial laboratory investigations revealed hypo-osmolar hyponatraemia with serum sodium of 127 mmol/L (135–145 mmol/L) and serum osmolality of 264 mmol/kg (275–295 mmol/kg), with an inappropriately concentrated urine osmolality of 649 mmol/kg (50–1200 mmol/kg) consistent with syndrome of inappropriate anti-diuresis. A microcytic anaemia with biochemical iron deficiency was identified. Renal and hepatic function were otherwise unremarkable. Vitamin B12, serum folate, and viral hepatitis serology were unremarkable. There was no family history of porphyria identified.
HIV infection was diagnosed three years earlier with an initial cluster of differentiation 4 (CD4) cell count of 229 cells/µL. Ongoing adherence challenges were noted. Since commencing ART, her CD4 count has increased to 383 cells/µL, while the HIV viral load has risen from 18 200 copies/mL to 77 905 copies/mL. Given the neurovisceral symptoms, tachycardia and hyponatraemia, acute porphyria was suspected. Biochemical porphyrin analyses are summarized in Table 1. A light protected spot urine sample screened strongly positive for porphobilinogen using the Hoesch test, a recommended rapid screening method for acute hepatic porphyrias, confirming an acute porphyric attack [7]. The urine PBG:creatinine ratio was markedly increased at 69 mmol/mmol (reference <1.5 mmol/mmol), in keeping with guideline-recommended quantitative confirmation of acute hepatic porphyria [8]. Total urinary porphyrins were elevated at 1431.6 nmol/mmol creatinine. Urine porphyrin fractionation by high performance liquid chromatography demonstrated predominantly raised uroporphyrin and heptacarboxylporphyrin [8]. The plasma fluorescence scan did not peak at 626–628 nm (arguing against variegate porphyria), and the faecal profile was not typical of hereditary coproporphyria.

Simultaneously, the faecal porphyrin profile demonstrated markedly increased uroporphyrin, heptacarboxylporphyrin, and isocoproporphyrin, pathognomic of PCT. A plasma fluorescence peak at approximately 618nm, although non-specific, further supported a PCT-like biochemical profile [8]. Given the mixed biochemical pattern, *HMBS* gene sequencing was performed. This confirmed a heterozygous NM_000190.4(HMBS):c.770dup p. (Glu258Glyfs*33) frameshift variant, classified as likely pathogenic, establishing a molecular diagnosis of AIP [9]. The PCT component was attributed to acquired hepatic UROD enzyme inhibition, likely triggered by HIV-related hepatic oxidative stress and cytochrome P450 induction in the absence of iron overload or viral hepatitis. *UROD* gene sequencing is not routinely available and therefore true dual porphyria could not be excluded.

Together, the findings demonstrate AIP with a PCT-like hepatic porphyrin pattern likely due to acquired UROD inhibition. During the acute attack, the patient was managed with non-porphyrinogenic opioid analgesia for pain and intravenous 5% dextrose to suppress hepatic 5-aminolevulinate synthase 1 (ALAS1) transcription. Heme arginate was administered to provide negative feedback on ALAS1, rapidly reducing

production of neurotoxic intermediates (ALA and PBG) and shortening the duration of the attack.

During each episode, her clinical condition improved following heme therapy, pain control, and correction of electrolyte abnormalities.

Because the AIP attacks were thought to be related to the menstrual cycle, a gonadotropin-releasing hormone (GnRH) analogue was initiated for chronic management. GnRH analogues suppress ovarian steroidogenesis by downregulating pituitary GnRH receptors, resulting in suppression of cyclic progesterone surges, which induces ALAS1. In hormonally sensitive AIP, this stabilises hormonal cycling, reduces the high-risk luteal window and helps prevent recurrent neurovisceral attacks.

Table 1: Biochemical investigations of urine, blood, and faeces in the evaluation of porphyria.

Test	Result	Reference interval	Units
Urine porphyria investigations			
Total porphyrins	1431.6 (H)	<300	nmol/mmol creatinine
Uroporphyrin	1120.2 (H)	<40	nmol/mmol creatinine
Heptacarboxylporphyrin	153.5 (H)	<10	nmol/mmol creatinine
Hexacarboxylporphyrin	50.4 (H)	<5	nmol/mmol creatinine
Pentacarboxylporphyrin	67.8 (H)	<5	nmol/mmol creatinine
Coproporphyrin I	21.3	<50	nmol/mmol creatinine
Coproporphyrin III	18.4	<110	nmol/mmol creatinine
Blood porphyria investigations			
Plasma porphyrins	Positive	Negative	
Plasma fluorescence scan	Emission peak at 618 nm		
Faecal porphyria investigations			
Total porphyrins	1840 (H)	<200	nmol/g dry mass
Coproporphyrin I	310 (H)	<200	nmol/g dry mass
Coproporphyrin III	1150 (H)	<200	nmol/g dry mass
Protoporphyrin	380 (H)	<150	nmol/g dry mass
Isocoproporphyrins	Present	Absent	

H = above the reference interval

Figure 1: Heme biosynthesis pathway and relevant enzyme defects [10].

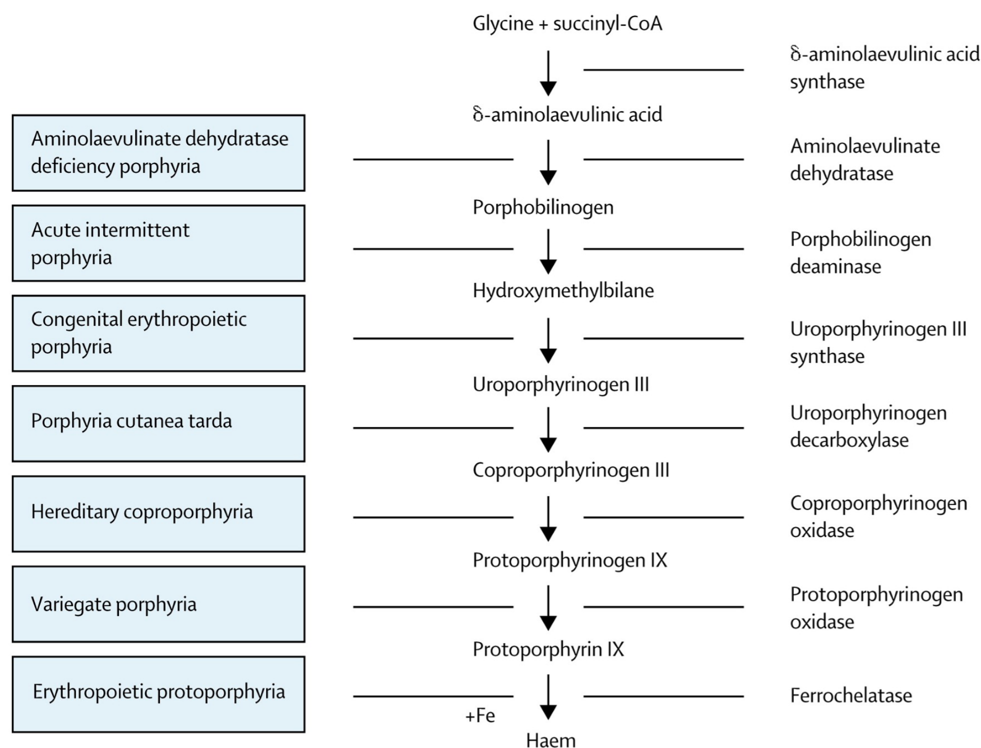


Figure 1 illustrates the heme biosynthesis pathway. AIP results from reduced HMBS activity, leading to accumulation of ALA and PBG and acute neurovisceral attacks. Inhibition of UROD activity produces a PCT-like porphyrin pattern with increased uroporphyrins and isocoproporphyrins, as observed in this patient.

Discussion

The markedly elevated urinary porphobilinogen during the acute neurovisceral episodes, together with the absence of a plasma fluorescence peak at 623–626 nm and the lack of increased faecal coproporphyrins, strongly supported an acute hepatic porphyria consistent with AIP rather than Variegate Porphyria or Hereditary Coproporphyria [1,8]. Identification of a pathogenic variant in the HMBS gene confirmed the diagnosis of AIP in this patient [9].

At the same time, the faecal porphyrin profile showed a dominant uroporphyrin and heptacarboxylporphyrin pattern with clear formation of isocoproporphyrins. This profile is characteristic of PCT or a PCT-like hepatic porphyrin pattern and reflects decreased activity of the enzyme uroporphyrinogen decarboxylase [6]. PCT arises when hepatic UROD activity falls below roughly 20%, often due to oxidative conversion of uroporphyrinogen into inhibitor porphyrinogens under conditions of hepatic stress [6,8].

The simultaneous presence of an AIP profile and a PCT-like biochemical pattern raises two possible explanations. The first is true dual porphyria caused by inherited defects in both the HMBS and UROD genes. Figure 1 illustrates the heme biosynthesis pathway, highlighting the enzymatic blocks relevant to AIP and the downstream porphyrin accumulation observed in this case [8]. The second is a functional dual pattern in which a single inherited defect in the HMBS gene combines with acquired suppression of uroporphyrinogen decarboxylase due to hepatic stress.

Despite the lack of genetic sequencing, clinical evidence suggests this case represents a secondary PCT-like pattern rather than a primary defect. Notably, the patient exhibited iron deficiency, which contradicts the classic PCT phenotype of iron overload. However, localized hepatic oxidative stress can impair UROD enzyme function regardless of systemic iron status. Furthermore, iron deficiency may exacerbate the underlying AIP by limiting substrate availability for ferrochelatase, thereby further reducing the regulatory heme pool and upregulating ALAS1 activity [11]. We propose that chronic HIV infection served as the primary driver; HIV-induced oxidative stress and mitochondrial impairment likely facilitated the formation of oxidized porphyrinogens [5,6]. These molecules inhibit UROD activity, effectively producing a biochemical phenotype of PCT [5].

The patient was receiving a tenofovir disoproxil fumarate, lamivudine, and dolutegravir–based antiretroviral regimen, with documented adherence challenges and virological non-suppression. While this regimen is generally well tolerated, both chronic HIV infection and antiretroviral exposure have been associated with mitochondrial dysfunction, and hepatic oxidative stress, even in the absence of overt biochemical liver injury. In this context, subclinical hepatic metabolic stress may plausibly contribute to secondary suppression of uroporphyrinogen decarboxylase activity, providing a mechanistic explanation for the observed PCT-like biochemical profile despite normal liver enzymes and the absence of classical iron overload.

Pathophysiologically, both HIV infection and antiretroviral therapy can promote hepatic oxidative stress and metabolic

dysfunction that may secondarily inhibit UROD [5,6,12]. Chronic HIV infection has been associated with increased generation of reactive oxygen species, altered iron handling, and CYP450 induction, all of which favour oxidative conversion of uroporphyrinogen to inhibitory porphyrinogens within hepatocytes [5,6,11]. Several ART agents may further contribute through mitochondrial toxicity, steatotic liver change and disruption of hepatic redox pathways, even when conventional liver biochemistry is normal [5,12]. When hepatic UROD activity is sufficiently reduced, a PCT-like porphyrin pattern emerges with excess uroporphyrins and isocoproporphyrins, as seen in this case [4,6,8]. In this context, poor virological control may lower the threshold for secondary UROD inhibition and modify the biochemical expression of underlying AIP [5,13].

True dual porphyria was considered but deemed less likely due to the absence of persistent biochemical features across multiple sample types and the lack of confirmatory genetic evidence [9].

Drug-induced pseudoporphyria and antiretroviral-associated photosensitivity were also considered in the differential diagnosis [5]. However, pseudoporphyria is typically characterised by cutaneous blistering in the absence of porphyrin over-production, whereas this patient demonstrated marked increases in urinary uroporphyrin and heptacarboxylporphyrin together with faecal isocoproporphyrins, which are regarded as pathognomonic for hepatic UROD inhibition [4,6,7]. Furthermore, the patient was receiving a dolutegravir-based regimen, and antiretroviral-related photosensitivity is both uncommon with this regimen and not associated with the porphyrin excretion pattern observed [5]. These findings support a true PCT-like biochemical phenotype rather than drug-induced pseudoporphyria.

From the perspective of laboratory medicine, this case shows that accurate porphyria diagnosis cannot rely on a single specimen type [7,10,14]. Interpretation depends on integrating information from several sample types, including urine for early pathway precursors, plasma for fluorescence patterns that separate acute porphyrias, and faeces for downstream porphyrins [8,10,14]. A PCT- type faecal porphyrin pattern should therefore be regarded as evidence of reduced hepatic uroporphyrinogen decarboxylase activity even in the absence of skin findings [4,6].

This case also illustrates the value of careful attention to pre analytical factors, particularly in settings with limited resources. Protection of samples from light, appropriate timing of specimen collection during clinical attacks and thoughtful review of concurrent medications, iron status and hormonal influences all contribute to accurate interpretation of porphyrin studies.

Patient Outcome

At present, the patient continues to experience acute porphyria attacks, though their frequency and severity have decreased since the initiation of GnRH analogue therapy. She remains under the care of the endocrine outpatient department, where management is focused on maintaining virological suppression and minimizing hormonal triggers. Long-term follow-up aims to determine if the secondary PCT-like biochemical profile is transient or persistent; this will be assessed through serial porphyrin studies performed during periods of clinical stability.

Limitations

UROD gene sequencing was not available. Although the biochemical pattern strongly suggests secondary hepatic UROD inhibition rather than hereditary PCT, the absence of UROD sequencing means that true dual porphyria (co-existence of pathogenic HMBS and UROD variants) cannot be completely excluded.

Secondly, although we propose that HIV-mediated oxidative stress was the primary driver of the PCT-like profile, the lack of independent assessment for liver fibrosis or steatosis is a limitation. Without this data, we cannot fully exclude the presence of underlying structural liver disease as a synergistic factor in impairing UROD function.

A further limitation is that repeat porphyrin studies were only performed during subsequent acute attacks, rather than in the period of clinical stability immediately following therapy. While these follow-up results confirmed the persistence of the underlying AIP, the lack of testing during a stable, asymptomatic interval precludes an assessment of whether the secondary PCT-like biochemical profile resolved with improved HIV infection control.

Conclusion

This patient's biochemical porphyria pattern is most likely due to a primary HMBS deficiency causing acute attacks, while oxidative stress from uncontrolled HIV infection and hormone-driven ALAS1 induction probably caused secondary UROD inhibition.

This case shows the need for planned and consistent monitoring in HIV-positive patients with porphyria, especially when the viral load is not well controlled. The key goals are to achieve viral suppression, assess liver function and iron status at intervals, and avoid medicines that are hepatotoxic or porphyrin-inducing. To help distinguish ongoing secondary UROD inhibition from true PCT, clinicians should consider serial urine, stool and plasma porphyrin profiles. It is also important to address other causes of oxidative stress, such as alcohol use, oestrogen exposure, or viral hepatitis.

Learning points

- In settings with limited resources, reliable diagnosis depends on strict sample handling and careful clinical correlation because repeat sampling, full porphyrin profiles, and genetic confirmation may not be readily accessible.
- Isocoproporphyrins detected in faeces are a defining marker of hepatic UROD inhibition and confirm a PCT-type biochemical pattern.
- Hepatic oxidative stress, mitochondrial dysfunction, and cytochrome P450 induction associated with HIV infection and ART can secondarily suppress UROD activity, producing a PCT-like phenotype despite normal UROD genotype.
- PCT-like profiles occur when hepatic UROD activity falls below ~20%, typically due to oxidative conversion of uroporphyrinogen into inhibitory porphyrinogens.
- Progesterone-driven ALAS1 up-regulation during the luteal phase increases precursor production and can trigger cyclical neurovisceral attacks in hormonally sensitive AIP.

Author Disclosures and contributions

Ethical Clearance

Written informed consent was obtained from patient for publication of this case. Ethical clearance was obtained from the Medical Human Research Ethics Committee, University of the Witwatersrand - clearance certificate no. M250766.

Funding

No specific funding was received for this case report.

Author contributions

Conceptualization: KM Masemola, KR Masemola, M Dintshi, T Pillay.
Investigation and data curation: KM Masemola, T Pillay.
Supervision: M Dintshi, T Pillay.
Writing – original draft: KM Masemola.
Writing – review and editing: KR Masemola, M Dintshi, T Pillay.
All authors approved the final version of the manuscript.

AI disclosure

The manuscript was edited using ChatGPT (OpenAI) to improve wording and clarity. No AI tools were used to generate, change, or analyze any clinical data, laboratory results, figures, or interpretations. All scientific content and conclusions reflect the authors' own work and judgement.

Competing interests

The authors declare that they have no competing interests.

References

1. Phillips JD. Heme biosynthesis and the porphyrias. Mol Genet Metab. 2019;128(3):164-177. doi: 10.1016/j.ymgme.2019.04.008. Epub 2019 Apr 22. PMID: 31326287; PMCID: PMC7252266.
2. Balwani M, Desnick RJ. The porphyrias: advances in diagnosis and treatment. Hematology Am Soc Hematol Educ Program. 2012;2012(1):19-27. doi:10.1182/asheducation-2012.1.19.
3. Puy H, Gouya L, Deybach JC. Porphyrias. Lancet. 2010;375(9718):924-937. doi:10.1016/S0140-6736(09)61925-5.
4. Leaf RK, Dickey AK. Porphyria cutanea tarda: a unique iron-related disorder. Hematology Am Soc Hematol Educ Program. 2024;2024(1):450-456. doi: 10.1182/hematology.2024000664. PMID: 39644053; PMCID: PMC11665602.
5. Isaacs T, Lehloenya R. HIV-associated photodermatitis in African populations. Front Allergy. 2023;4:1159387. doi: 10.3389/falgy.2023.1159387. PMID: 37216149; PMCID: PMC10192905.
6. Singal AK. Porphyria cutanea tarda: Recent update. Mol Genet Metab. 2019;128(3):271-281. doi: 10.1016/j.ymgme.2019.01.004. Epub 2019 Jan 18. PMID: 30683557.
7. Di Pierro E, De Canio M, Mercadante R, Savino M, Granata F, Tavazzi D, Nicolli AM, Trevisan A, Marchini S, Fustinoni S. Laboratory Diagnosis of Porphyria. Diagnostics (Basel). 2021;11(8):1343. doi: 10.3390/diagnostics11081343. PMID: 34441276; PMCID: PMC8391404.
8. Wang B, Bonkovsky HL, Lim JK, Balwani M. AGA Clinical Practice Update on Diagnosis and Management of Acute Hepatic Porphyrias: Expert Review. Gastroenterology. 2023;164(3):484-491. doi: 10.1053/j.gastro.2022.11.034. Epub 2023 Jan 13. PMID: 36642627; PMCID: PMC10335308.
9. Sardh E, Barbaro M. Acute Intermittent Porphyria. 2005 Sep 27 [Updated 2024 Feb 8]. In: Adam MP, Bick S, Mirzaa GM, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2025. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1193
10. Kauppinen R. Porphyrias. Lancet. 2005;365(9455):241-252. doi:10.1016/S0140-6736(05)17744-7.
11. Camaschella C, Nai A, Silvestri L. Iron metabolism and iron disorders revisited in the hepcidin era. Haematologica. 2020;105(2):260-272. doi: 10.3324/haematol.2019.232124. PMID: 31949017; PMCID: PMC7012465.
12. Wang B, Rudnick S, Cengia B, Bonkovsky HL. Acute Hepatic Porphyrias: Review and Recent Progress. Hepatol Commun. 2018;3(2):193-206. doi: 10.1002/hep4.1297. PMID: 30766957; PMCID: PMC6357830.
13. Stein PE, Badminton MN, Rees DC. Update review of the acute porphyrias. Br J Haematol. 2017;176(4):527-538. doi:10.1111/bjh.14459.
14. Anderson KE, Bloomer JR, Bonkovsky HL, Kushner JP, Pierach CA, Pimstone NR, et al. Recommendations for the diagnosis and treatment of the acute porphyrias. Ann Intern Med. 2005;142(6):439-450. doi:10.7326/0003-4819-142-6-200503150-00010.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

Monoclonal Gammopathy with Double M-bands Mimicking Biclonal Gammopathy: a case series

Dharmendra Kumar[1], Koushik Biswas[1], Kanu Tiwari[2], Mala Mahto[2*]

[1]Department of Biochemistry, All India Institute of Medical Sciences, Raebareli, U.P., India
[2]Department of Biochemistry, All India Institute of Medical Sciences, Patna, Bihar, India

Article Info

Corresponding Author:
Mala Mahto Professor
Department of Biochemistry
All India Institute of Medical Sciences, Patna, Bihar, India, 801507
E-mail: dr.malamahto@gmail.com
ORCID ID: 0000-0003-0445-6972

Keywords

monoclonal gammopathy, multiple myeloma, blood protein electrophoresis, immunoglobulins

Abstract

Background: Monoclonal gammopathies are characterised by the presence of a serum or urine monoclonal protein (M-protein), typically detected as a single narrow band on serum protein electrophoresis (SPEP). Rarely, double M-bands may appear, raising the suspicion of biclonal gammopathy. Distinguishing true biclonality from paraprotein polymerization is crucial to prevent diagnostic misclassification and inform appropriate clinical decisions.

Case: We report a case series of four multiple myeloma patients with twin M-bands who presented to XXXXX between March 2023 and March 2024. All four patients exhibited two discrete M-bands on SPEP, initially suggestive of biclonal gammopathy. However, immunofixation electrophoresis confirmed that both bands in each case belonged to the same immunoglobulin isotype and light chain, establishing a monoclonal origin. Treatment of samples with β-mercaptoethanol resulted in the collapse of double bands into a single monoclonal band, confirming paraprotein polymerization. IgA and IgM paraproteins were most frequently implicated in dual-band appearance.

Conclusion: The presence of double M-band on serum protein electrophoresis does not necessarily indicate biclonal gammopathy. Depending on SPEP results alone may lead to misclassification, and confirmatory techniques such as IFE and reducing agent studies are essential to establish the correct diagnosis.

Introduction

Monoclonal gammopathy (MG) is the presence of serum and/or urine monoclonal protein (M-protein), which is caused by the overproduction of the immunoglobulins (Igs) by a population of B-lymphocytes and/or plasma cells. This is primarily observed in Waldenstrom's macroglobulinemia, multiple myeloma (MM), and monoclonal gammopathy of unknown significance (MGUS). However, primary amyloidosis, cryoglobulinemia, lymphoproliferative diseases, solitary plasmacytoma, hepatitis C infection, POEMS syndrome, and scleromyxedema can also result in monoclonal deposition of other proteins [1]. In serum protein electrophoresis (SPEP), Longsworth et al. [2] initially reported a tall, narrow-based spike that Moore et al. [3] later referred to as the "M" component/peak. When doing SPEP, polyclonal Igs show up as a broad-based band on agarose gel, while monoclonal Igs show up as a single, discrete and intense band on agarose gel and a distinct peak in the densitometer tracing. Depending on the type of Igs produced, this band or peak may also be seen in the gamma, beta-gamma, or beta-region. Following the discovery of the "M" band or peak on SPEP, immunofixation electrophoresis (IFE) is used to definitively identify monoclonal proteins and characterize the heavy chain (HC) and light chain (LC) [4]. IFE is more sensitive than SPEP in identifying traces of "M" protein in plasmacytoma, primary amyloidosis, or treated cases of multiple myeloma/macroglobulinemia, where SPEP does not show a monoclonal band. Additionally, it helps differentiate between a monoclonal and polyclonal increase in Igs [5]. Monoclonal immunoglobulins (Igs) are produced by the growth of neoplastic plasma cells, but an increase in polyclonal Igs is usually linked to reactive or inflammatory events. Moreover, if there is a strong clinical suspicion of multiple myeloma or a similar condition in cases with normal SPEP, IFE is warranted to confirm the diagnosis [6].

A subgroup of patients with monoclonal gammopathies may have serum and/or urine containing several M-proteins. This group of gammopathies is known as "double gammopathies." In existing literature, the terms "biclonal gammopathy" and "double gammopathy" have been used interchangeably. Double gammopathy can be depicted in two different ways, wherein two different M bands or peaks on SPEP or a single M band on SPEP further resolves into two distinct bands on IFE. The variant of HC and LC found determines the further subdivision of double gammopathies. In these situations, IFE may reveal whether the monoclonal proteins differ in both HC and LC isotype (e.g IgA kappa and IgG lambda) which is a true biclonal gammopathy. In some cases the M proteins share the same HC with different LCs or different HCs with same LCs which is termed double gammopathies (IgM kappa and IgM lambda or IgA kappa and IgG kappa) [7]. Multiple myeloma, several lymphoproliferative diseases, and MGUS have been linked to biclonal/double gammopathies.

Biclonal/Double gammopathies can be picked up at initial presentation or after the original diagnosis has been established and the disease has progressed. The eruption of "distinct plasma cell clones that are two or more in number" producing multiple monoclonal proteins can lead to biclonal/double gammopathies[7,9]. At a later stage of disease pathogenesis, when "clonally related" plasma cells undergo class switching, it can also result in biclonal/double gammopathies. Throughout the course of the disease or after treatment, the proportion of "double" M-proteins may increase or decrease in a way that is consistent or inconsistent. The causative mechanism of this process is unclear [8]. Biclonal/Double gammopathies are reported to occur in 2-4 per cent of different gammopathy cohorts in the published literature [8,10-14]. The most prevalent combination of HC isotype, according to Kyle et al. [10], is IgG–IgA. Riddell et al. [8] reported that the most prevalent HC isotype combination was IgG–IgM. Two other studies reported that the most prevalent combination is IgG–IgG [13,14]. These findings differ from one another for a variety of reasons, including variation in racial or population distribution, variation in the sensitivity of the detection technique or selection bias in the process of case selection. Most frequently reported light chains were the kappa light chain [8,11,14]. This case series presents four multiple myeloma patients with twin M-bands on SPE who presented to XXXXX between March 2023 and March 2024. All patients had provided informed consent to the publication of their anonymised data.

Results

All four patients demonstrated double M-bands on SPEP or IFE. In all, the presence of two distinct bands initially suggested biclonal gammopathy. On subsequent IFE, both bands in each except the last case (case 4) were found to belong to the same immunoglobulin heavy and light chain type, confirming a monoclonal gammopathy with pseudo dual-band appearance. To further evaluate the possibility of polymerization or aggregation, serum samples were treated with the reducing agent β-mercaptoethanol before electrophoresis. After reduction of the sulfide bonds, the double M-bands collapsed into a single monoclonal band, supporting the interpretation that the apparent biclonal pattern was due to paraprotein polymerization giving rise to two molecules with different molecular weights (partly polymerized and partly unpolymerized) rather than the presence of two independent clones. The description of the cases is presented in Table 1. SPEP and IFE of the four cases are presented in Figures 1 to 4. The SPE and IFE were performed on a gel-based electrophoresis machine (Hydrasys 2, Sebia, France).

Table 1: Clinical features and laboratory findings of cases.

Case	Age/Sex	Clinical findings	Laboratory findings	Bone marrow findings	SPEP finding	IFE finding	Diagnosis
1	57y/M	Recurrent chest pain, shortness of breath, weakness, O/E- pallor and B/L basal crepts	Hb- 46 g/L, TP /ALB/GLOB-114.1/22.8/84.3 g/L, A/G ratio-0.35, Cr-193.6 µmol/L, Ca-2.24 mmol/L.	72% plasma cell (binucleate /multinucleate)	Two peak M band beta gamma region	IgA lambda	Plasma cell dyscrasia
2	56y/F	Fatigue, weakness, vomiting and anorexia k/c/o HTN and bronchial asthma	Hb- 68 g/L, TP /ALB/GLOB-56.7/28.7/28.0 g/L, A/G ratio-1.03, Cr-1096 µmol/L, Ca-2.45 mmol/L.	11% plasma cell (binucleate /trinucleate)	Two peak M band beta gamma region	IgG lambda	Plasma cell dyscrasia
3	75y/F	Fever, back pain, weakness O/E- pallor	Hb- 78 g/L, TP /ALB/GLOB-96.7/27.9/68.8 g/L, A/G ratio-0.41, Cr-107.0 µmol/L, Ca-2.77 mmol/L.	50% plasma cell (binucleate /trinucleate)	M band in beta gamma region	IgA kappa	Plasma cell dyscrasia
4	80y/F	Weakness, loose stool and abdominal pain, O/E- pallor	Hb-79 g/L, TP/ ALB/GLOB-87.1/33.0/54.1 g/L, A/G ratio-0.61, Cr- 158.3 µmol/L, Ca-2.22 mmol/L.	40% plasma cell (multinucleate)	M band in beta gamma region	True Biclonal gammopathy (IgG Kappa and IgA lambda)	Plasma cell dyscrasia

SPEP: serum protein electrophoresis, IFE: immunofixation electrophoresis, M: Male, F: Female, Hb: Hemoglobin, TP: Total Protein, ABL: Albumin, GLOB: Globulin, A/G: Albumin/Globulin, Cr: Creatinine, Ca: Calcium, O/E: on examination, HTN: hypertension, B/L: bilateral, k/c/o: known case of, IgA: Immunoglobulin A, IgG: Immunoglobulin G, M band: monoclonal band, g/dl: grams per liter; µmol/L: micromoles per liter; mmol/L: millimoles per liter

Figure 1: Serum protein electrophoresis and immunofixation of Case 1 before and after β-mercaptoethanol treatment.

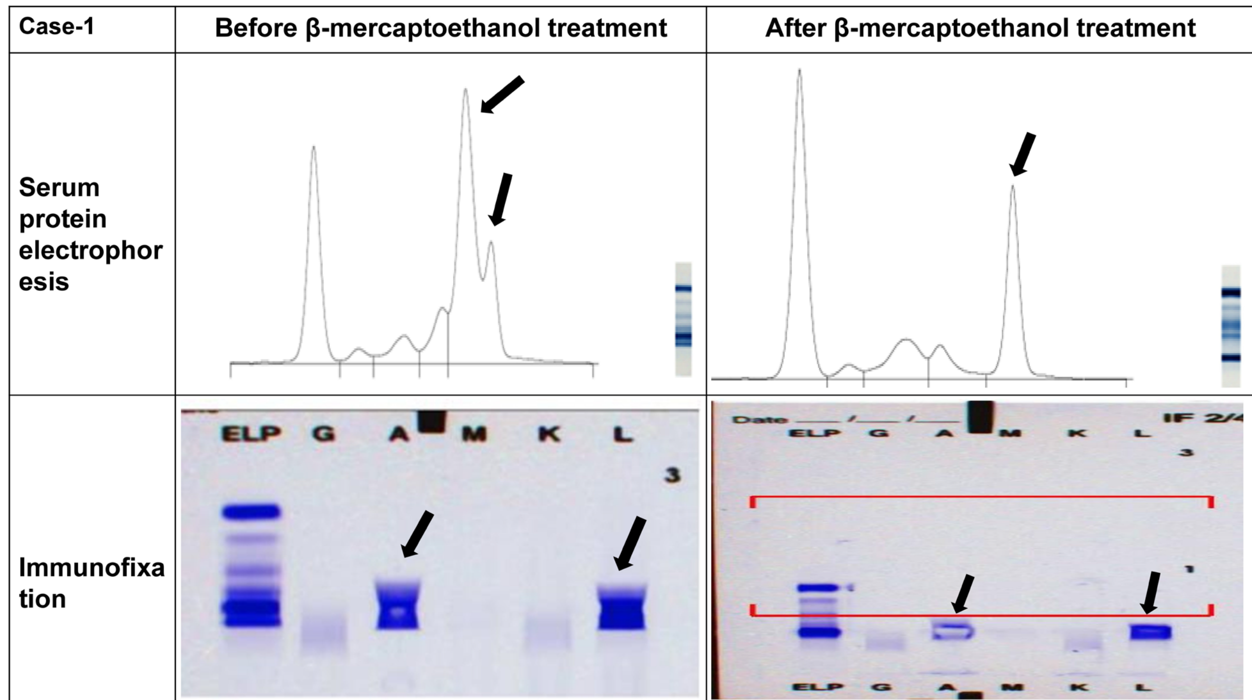


Figure 2: Serum protein electrophoresis and immunofixation of case 2 before and after mercaptoethanol treatment.

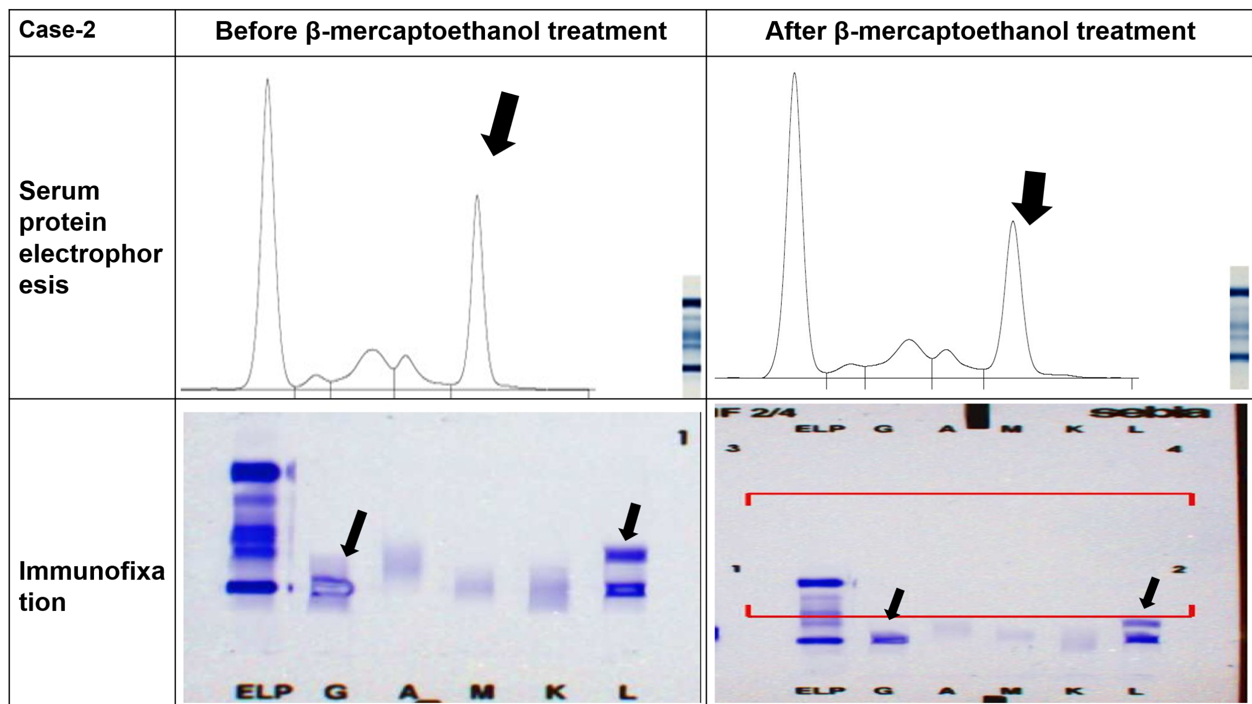


Figure 3: Serum protein electrophoresis and immunofixation of Case 3 before and after β-mercaptoethanol treatment.

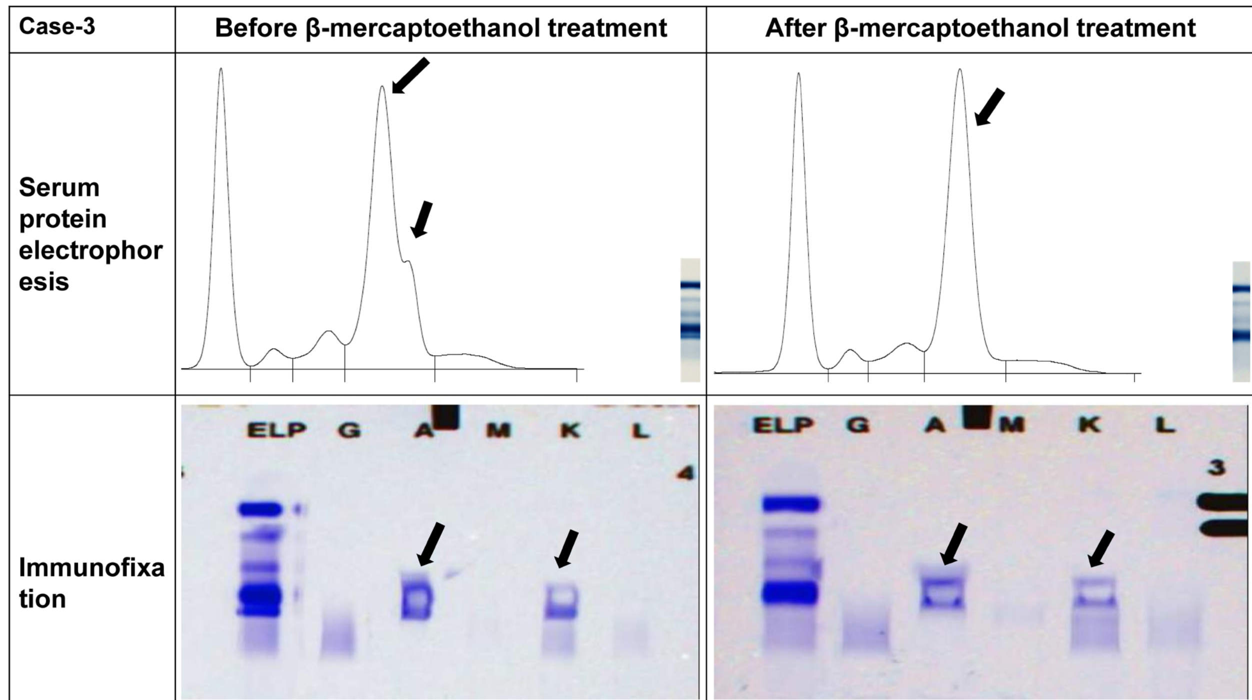
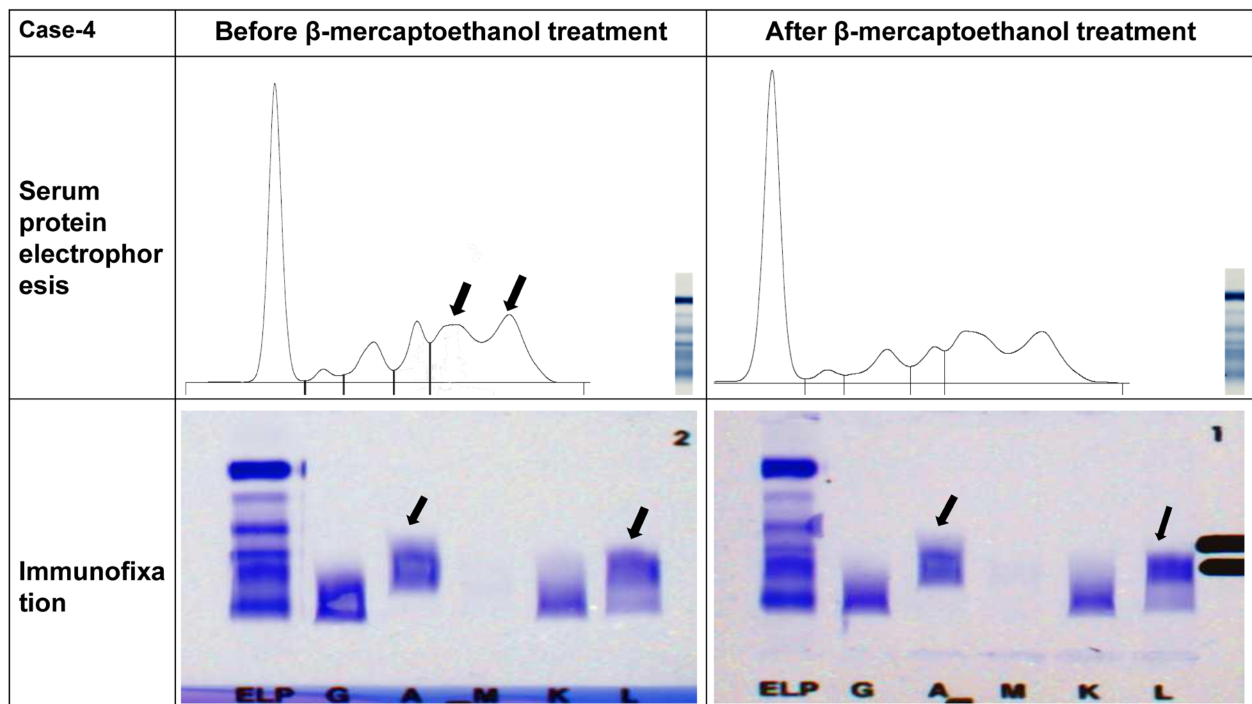


Figure 4: Serum protein electrophoresis and immunofixation of Case 4 before and after β-mercaptoethanol treatment.



Discussion

Multiple myeloma is the second most prevalent hematological malignancy, contributing to nearly 2% of cancer-related mortality worldwide [15]. SPEP is an essential diagnostic tool in the initial laboratory evaluation and characterization of monoclonal proteins. Monoclonal immunoglobulins usually appear on SPEP as a sharp band within the gamma-globulin region. However, due to their unique physicochemical properties, monoclonal IgM andespecially IgA may migrate into the beta region [16]. Often, IgA-type M proteins exhibit unusual and atypical electrophoretic patterns, which can complicate interpretation. These include a tendency for more anodal migration toward the β-region due to their lower isoelectric point. IgA has a quaternary structure that permits polymerization, with or without subsequent light chain (LC) production, due to the sequestration of LC epitopes, thereby hindering the reaction between LC epitopes and their antibodies.

This produces a condition termed 'IgA with no apparent light chain attached', an apparent lack of detectable light chains on IFE. Another recognized cause of pseudo-biclonal gammopathy is the tendency of immunoglobulins to polymerize and form aggregates. IgA and IgM are the immunoglobulins most commonly involved in polymer formation. In such cases, the polymerized form has a different molecular weight as compared to the unpolymerized form, leading to different localization on SPE due to a difference in relative migration. A possible interference from a fibrinogen band leads to a common misinterpretation of biclonal gammopathy due to the presence of an additional band in the gamma region, in addition to a single M band in the case of true monoclonal gammopathy. Fibrinogen is noted in the gamma region on SPE when a plasma sample is used instead of serum due to probable contamination with EDTA/heparin or other anticoagulants. Subsequent IFE rules out the probability of biclonal gammopathy as it does not employ antisera targeted to fibrinogen. Some cases with more than one clone of M protein are also noted in patients with remission of monoclonal gammopathy following treatment or during progression from less aggressive to more aggressive forms and renal failure [17,18].

In this study, on SPE, cases 1 and 2 revealed a bifurcation of the sharp band obtained in the beta-gamma region, hinting at the presence of a biclonal gammopathy. IFE revealed IgA lambda and IgA kappa monoclonal gammopathy, respectively. The pseudo-biclonal pattern observed in cases 1 and 2 (Figures 1 and 2) on SPE was most likely due to polymerization of IgA, leading to the appearance of an additional M-band near the beta region, which subsequently resolved as IgA Lambda upon IFE. Upon BME treatment, polymers collapsed into monomers, resulting in a single band if it is a true monoclonal protein representing a single clone (like polymerized IgA).

In Case 3, a single monoclonal band was observed in the gamma region on SPE. However, the biclonal pattern resolved into IgG-Lambda-Lambda (Light Chain isotype-matched). This finding is probably because of the uncoordinated production of excess free light chains in serum, which can migrate faster due to lower molecular weight and produce an additional band in the lambda light chain lane. In Case 4, two peaks were observed in the gamma region on SPE. On IFE, two separate clones were identified, namely IgG kappa and IgA lambda, indicating a biclonal gammopathy. Further on treatment with BME, the same bands persisted, indicating a true biclonal gammopathy of IgG Kappa and IgA lambda type.

True biclonal gammopathy occurs only in clonally related plasma cells, of which IgG-IgA, IgG-IgM and IgG-IgG are reported in various literature [12,13]. These are caused by two main mechanisms, either heavy chain or light chain isotype matching. Distinguishing true biclonality from polymerization-related artefacts is essential, as misinterpretation can result in overestimation of disease severity, unwarranted investigations, or inappropriate treatment decisions. In genuine biclonal gammopathy, two discrete monoclonal proteins are typically produced by separate plasma cell clones. On rare occasions, however, a single clone may generate two M-proteins through mechanisms such as class-switch recombination or dual isotypeexpression events, that, while uncommon, are documented in the literature [19]. Conversely, particularly in the case of IgA, a single monoclonal protein may undergo polymerization and produce multiple migrating bands on electrophoresis, thereby simulating biclonality [20]. Further, a pentameric IgM may dissociate into 7S subunits and present on electrophoresis as one or more additional monoclonal bands. These bands are generally attributable to alternate molecular forms of a single monoclonal protein produced by the same plasma cell clone [1]. In our cases, this mechanism was substantiated, as treatment with β-mercaptoethanol unmasked a single IgA lambda component in three cases of false biclonal gammopathy.

Chemical reduction with β-mercaptoethanol serves as a useful diagnostic approach in such contexts. By cleaving the disulfide bonds that join immunoglobulin monomers into dimers or higher-order polymers, β-mercaptoethanol dissociates these complexes into their monomeric form, which subsequently appears as a single, uniform peak on SPEP [17]. In our cases, reduction with β-mercaptoethanol resulted in the convergence of the two discrete bands into a single band, thereby substantiating IgA polymerization as the underlying mechanism rather than true biclonality.

Beyond mere detection, IgA polymerization carries important implications for disease monitoring. Polymeric IgA can be erroneously interpreted as two distinct M-proteins, resulting in an overestimation of disease burden. Evidence suggests that patients with polymeric IgA often demonstrate apparently higher levels of M-protein compared to those with monomeric IgA [7,17]. Additionally, IgA polymers have been linked to clinical complications such as hyperviscosity syndrome and laboratory artefacts, including spuriously raised calcium levels due to calcium binding by IgA complexes [16,20]. While our cases did not develop hyperviscosity syndrome or hypercalcemia, these potential risks highlight the clinical relevance of identifying IgA polymerization. Although IFE is considered to be gold standard for paraprotein characterization, it may not single-handedly suffice in distinguishing co-migrating polymers and may sometimes lead to inconclusive results [11]. In our cases, IFE was performed both before and after pretreatment with β-mercaptoethanol to aid in clarification.

Conclusion

This case series highlights that the presence of double M-band on serum protein electrophoresis does not necessarily indicate biclonal gammopathy. In all our patients, immunofixation electrophoresis confirmed a single monoclonal protein, and in selected cases, the use of reducing agents such as β-mercaptoethanol further demonstrated that the dual bands were due to polymerization of the same paraprotein rather than two independent clones. The observation was most frequently associated with IgA paraproteins, which are inherently prone to form polymeric structures and exhibit atypical migration patterns. Therefore, reliance on SPEP alone may lead to misclassification, and confirmatory techniques such as IFE and reducing agent studies are essential to establish the correct diagnosis. Accurate interpretation is crucial for appropriate patient management and to avoid over-reporting of biclonal gammopathy.

Conflict of Interest

None.

Ethical Approval

Ethical clearance is waived off for case series from our institute. The study was done in accordance with the Declaration of Helsinki.

Author's contributions

All authors fulfil the criteria for authorship as set out by the ICMJE and as recommended by the Committee on Publication Ethics (COPE).

Funding

None.

Data availability

All data related to the case series has been submitted along with the manuscript.

All reagents and instrument for running SPE and IFE have been provided by SEBIA, France.

All Clinical chemistry assays have been provided by Beckman Coulter, USA.

References

1. Katzmann J, Kyle RA, Lust J, Snyder M, Dispenzieri
2. Immunoglobulins and laboratory recognition of monoclonal proteins. In Neoplastic Diseases of the Blood. Springer New York. 2013. p. 565-588. https://doi.org/10.1007/978-1-4614-3764-2_29
3. Longsworth LG, Shedlovsky T, Macinnes DA.
4. ELECTROPHORETIC PATTERNS OF NORMAL AND PATHOLOGICAL HUMAN BLOOD SERUM AND PLASMA. J Exp Med. 1939;70(4):399-413. https://doi.org/10.1084/jem.70.4.399.
5. Moore DH, Kabat EA, Gutman AB. BENCE-JONES PROTEINEMIA IN MULTIPLE MYELOMA. J Clin Invest. 1943 ;22(1):67-75. https://doi.org/10.1172/JCI101370.
6. Zhu S, Li W, Lin M, Li T. Serum Protein Electrophoresis and Immunofixation Electrophoresis Detection in Multiple Myeloma. J Coll Physicians Surg Pak. 2021 ;31(7):864-867. https://doi.org/10.29271/jcpsp.2021.07.864.
7. Kuriakose E, Narayanan Unni Cheppayil S, Kuzhikandathil Narayanan S, Vasudevan A. A Study on Free Light Chain Assay and Serum Immunofixation Electrophoresis for the Diagnosis of Monoclonal Gammopathies. Indian J Clin Biochem. 2019;34(1):76-81. https://doi.org/10.1007/s12291-017-0718-6.
8. Deveci K, Korkmaz S, Sancakdar E, Acibucu D, Alkan F, Terzi H. The evaluation of serum protein and serum immunofixation electrophoresis results in patients with monoclonal and polyclonal gammopathy: A single center experience. Int J Blood Res Disord. 2015;2(2):1-4. https://doi.org/10.23937/2469-5696/1410017
9. Tschumper RC, Dispenzieri A, Abraham RS, Henderson KJ, Jelinek DF. Molecular analysis of immunoglobulin genes reveals frequent clonal relatedness in double monoclonal gammopathies. Blood Cancer J. 2013;3(4):e112. https://doi.org/10.1038/bcj.2013.12.
10. Riddell S, Traczyk Z, Paraskevas F, Israels LG. The double gammopathies. Clinical and immunological studies. Medicine (Baltimore). 1986;65(3):135-142. https://doi.org/10.1097/00005792-198605000-00001.
11. Bakkus MH, Schots R, Gomez La Fuente PB, Van Riet I, Thielemans K, De Waele M, Van Camp B. Clonally related IgAand IgE-secreting plasma cells in a myeloma patient. Eur J Haematol. 2000;65(5):348-355. https://doi.org/10.1034/j.1600-0609.2000.065005348.x.
12. Kyle RA, Robinson RA, Katzmann JA. The clinical aspects of biclonal gammopathies. Review of 57 cases. Am J Med. 1981;71(6):999-1008. https://doi.org/10.1016/0002-9343(81)90326-0.
13. Kyle RA, Therneau TM, Rajkumar SV, Offord JR, Larson DR, Plevak MF, Melton LJ 3rd. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. N Engl J Med. 2002;346(8):564-569. https://doi.org/10.1056/NEJMoa01133202.
14. Kyle RA, Gertz MA, Witzig TE, Lust JA, Lacy MQ, Dispenzieri A, Fonseca R, Rajkumar SV, Offord JR, Larson DR, Plevak ME, Therneau TM, Greipp PR. Review of 1027 patients with newly diagnosed multiple myeloma. Mayo Clin Proc. 2003;78(1):21-33. https://doi.org/10.4065/78.1.21.
15. Guastafierro S, Ferrara MG, Sica A, Parascandola RR, Santangelo S, Falcone U. Serum double monoclonal components and hematological malignancies: only a casual association? Review of 34 cases. Leuk Res. 2012;36(10):1274-1277. https://doi.org/10.1016/j.leukres.2012.05.008.
16. García-García P, Enciso-Alvarez K, Diaz-Espada F, Vargas-Nuñez JA, Moraru M, Yebra-Bango M. Biclonal gammopathies: Retrospective study of 47 patients. Rev Clin Esp (Barc). 2015;215(1):18-24. English, Spanish. https://doi.org/10.1016/j.rce.2014.07.003.
17. Nayak BS, Ojar-Taylor N, St John S, Swann S, Thom J, Thomas B, Thomas L, Townsend D, Trotman S. Significance of Serum Protein Electrophoresis in the Detection of Multiple Myeloma: A Diagnostic Interpretation of Patients with Varied Immunoglobulins. Int J Prev Med. 2021;12:37. https://doi.org/10.4103/ijpvm.IJPVM_222_18.
18. Prisi S, Khurana V, Saijpaul R, Verma R, Chandra L, Koner BC. Unraveling the Possibilities of Monoclonal Protein Migration, Identification, and Characterization in SPEP on Capillary Zone Electrophoresis. J Lab Physicians. 2022;14(4):505-510. https://doi.org/10.1055/s-0042-1744242.
19. Srinivasan VK, Bhagat P, Bansal F, Chhabra S. Occurrence of Double Monoclonal Bands on Protein Electrophoresis: An Unusual Finding. Indian J Hematol Blood Transfus. 2016;32(Suppl 1):184-188. https://doi.org/10.1007/s12288-015-0622-2.
20. Bansal F, Bhagat P, Srinivasan VK, Chhabra S, Gupta P. Immunoglobulin A gammopathy on serum electrophoresis: A diagnostic conundrum. Indian J Pathol Microbiol. 2016;59(1):134-136. https://doi.org/doi: 10.4103/0377-4929.178245.

21. Sharma S, Gupta P, Aggarwal R, Malhotra P, Minz RW, Bansal F. Demystifying Biclonal Gammopathy: A Pathologist's Perspective. Lab Med. 2019;50(4):357-363. https://doi.org/10.1093/labmed/lmz006

22. Jain P, Choudhary R, Harith AK, Yadav C. Evaluation of Double M-Band on Serum Protein Electrophoresis Simulating Biclonal Gammopathy: A Case Report. Indian J Clin Biochem. 2022;37(2):247-249. https://doi.org/10.1007/s12291-020-00929-y.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

The Case of Negative Low Density Lipoprotein Cholesterol

Earnest JP Daniel[1,2], Sridevi Devaraj[1,2*]

[1]Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA
[2]Department of Pathology, Texas Children's Hospital, Houston, TX, USA

Article Info

Corresponding Author:
Sridevi Devaraj, Ph.D., DABCC, FADLM, FRSC
Department of Pathology and Immunology, Baylor College of Medicine, Texas Children's Hospital. 6621 Fannin Street, Houston, TX 77030
E-mail: sxdevara@texaschildrens.org

Keywords

LDL cholesterol, Hypertriglyceridemia, Low Cholesterol, Friedewald equation, Martin–Hopkins equation, Samson-NIH equation

Abstract

Accurate estimation of low-density lipoprotein cholesterol (LDL-C) is critical for lipid assessment and cardiovascular risk stratification, yet conventional calculation methods may fail in specific clinical contexts in pediatric population. The Friedewald equation, long considered standard, is unreliable in moderately elevated TG (TG ≥400 mg/dL) and in low total cholesterol (TC) scenarios. Alternative equations, including Martin–Hopkins and Sampson–NIH, have been developed to improve accuracy, but pediatric validation remains limited. We present two pediatric cases highlighting these challenges. Case 1 involved a 5-year-old male with perinatally acquired HIV/AIDS, low TC, and moderately elevated triglyceride (TG) <300 mg/dL. Friedewald-calculated LDL-C was negative, while direct beta-quantification revealed LDL-C of 36 mg/dL and Martin–Hopkins equation underestimated LDL-C by 70%. Case 2 involved a female with type 2 diabetes and high-risk B-cell acute lymphoblastic leukemia, with extreme hypertriglyceridemia >1000 mg/dL and elevated TC. In this case, both Friedewald and Martin–Hopkins calculations failed, whereas Sampson–NIH 2025 provided a calculable LDL-C of 26 mg/dL. These cases demonstrate that TG-to-TC ratios, rather than TG alone, predict calculation failure, and that alternative equations, particularly Sampson–NIH 2025, offer improved reliability. These findings are especially relevant for patients receiving cholesterol-lowering therapies, in whom low LDL-C and elevated TG may lead to significant underestimation of residual atherogenic risk. Further studies using comparative beta-quantification are needed to identify the most accurate LDL-C estimation methods across diverse pediatric populations and low-cholesterol states.

Introduction

Since its introduction in 1972, the Friedewald equation has served as a cornerstone of lipid assessment, providing a rapid and inexpensive estimate of low-density lipoprotein cholesterol (LDL-C) using the formula LDL-C = TC − HDL-C − TG/5 [1]. This calculation relies on several key assumptions, that very-low-density lipoprotein cholesterol (VLDL-C) represents approximately 20% of plasma triglycerides (TG) [1]. In most individuals, these assumptions hold sufficiently well to allow clinically acceptable LDL-C estimation. However, the Friedewald equation is known to fail when TG levels exceed 400 mg/dL (>4.52 mmol/L), at which point VLDL-C estimation becomes unreliable.

To address these limitations, several alternative equations have been developed [2-7]. The Martin–Hopkins equation, introduced in 2013 and validated primarily in adult populations, uses an adjustable TG: VLDL-C ratio derived from population-based strata using non-HDL-C and TG values [2]. This formula is a modified Friedewald equation with an adjustable factor, as LDL-C = TC − HDL-C − TG/adj factor. This approach provides more individualized VLDL-C estimation, improving LDL-C accuracy, particularly at low LDL-C concentrations (LDL-C <70 mg/dL or <1.8 mmol/L) and TG levels <400 mg/dL [2]. In 2023, an extended Martin-Hopkins equation was published with improved LDL-C accuracy in patients with moderate hypertriglyceridemia with TG even upto 800 mg/dL or 9.0 mmol/L [3]. A large review of 23 equation using >5000,000 patients suggested that this Martin–Hopkins formula remains the most accurate method for LDL-C calculation [4]. However, its clinical implementation by Lab Information Systems is limited by the need for a lookup table (~240 strata) to determine the adjustable factor [4-6].

More recently, in 2025, a modified version of the Sampson-NIH was published in Clinical Chemistry, where LDL-C is calculated as: LDL-C = non-HDL-C - (TG/8.37) - (TG x non-HDL-C/2640) + (TG2/ 17400) [7,8]. This modification builds upon the original equation described by Sampson et al in 2020, in which a multivariable regression-based equation was designed to more accurately model VLDL-C across a wider range of lipid profiles, including high triglycerides and low LDL-C, and was validated for triglyceride levels up to 800 mg/dL, outperforming theFriedewald equation in these setting [7,8]. The 2025 Sampson-NIH formula derived and validated in cohorts that included individuals with triglyceride levels exceeding 800 mg/d, reduces systematic underestimation at LDL-C levels <70 mg/dL (1.8 mmol/L) and showed closer agreement with beta-quantification, making it well suited for patients receiving modern lipid-lowering therapies [8,9]. In this report, we present two pediatric cases in which the Friedewald and Martin-Hopkins equations failed to provide a calculated LDL-C values, highlighting the limitations.

Case presentation

Case 1

A 5-year-old African American male with perinatally acquired HIV infection, currently with AIDS due to poor adherence to antiretroviral therapy, underwent routine lipid screening. Laboratory evaluation identified markedly low TC of 59 mg/dL, moderately elevated TG of 272 mg/dL, and low HDL-C of 11 mg/dL. LDL-C calculated using the Friedewald equation yielded a negative/incalculable value. Subsequent beta-quantification via ultracentrifugation revealed an LDL-C of 36 mg/dL, below the reference range but clearly quantifiable. Calculated LDL-C values using alternative equations demonstrated substantial negative bias: 16 mg/dL by the Martin–Hopkins equation, underestimating measured LDL-C by approximately 50–70% (Table 1)

Case 2

A female patient with type 2 diabetes mellitus and high risk B-cell acute lymphoblastic leukemia (HR B-ALL), undergoing chemotherapy complicated by severe hypertriglyceridemia and disc herniation, had extreme lipid derangements. Lipid studies revealed extreme hypertriglyceridemia (TG 1435 mg/dL), elevated total cholesterol (TC 221 mg/dL), HDL-C of 48 mg/dL, and non-HDL cholesterol of 173 mg/dL. Due to the extreme hypertriglyceridemia, LDL-C could not be calculated using the Friedewald or Martin–Hopkins equations; however, the Sampson–NIH 2025 equation yielded a calculated LDL-C of 26 mg/dL. (Table 1). No direct LDL-C assay was performed, and beta quantification was not available due to insufficient sample volume.

Table 1: Lipid parameters, TG/TC ratio, measured LDL-C (beta-quantification), and calculated LDL-C values using 3 different equations in both cases.

	TG	TC	HDL	Non-HDL (calculated)	TG/TC ratio	Beta-quant LDL	Calculated LDL (Friedewald equation)	Calculated LDL (Martin-Hopkins 2023)	Calculated LDL (modified Sampson-NIH 2025)
Ref Range	45-203 mg/dL	112-208 mg/dL	35-73 mg/dL	<145 mg/dL		60-140 mg/dL			
Case 1	272	59	11	48	4.6	36	-6 (Not calculable)	16	15
Case 2	1435	221	48	173	6.4	na	-114 (Not calculable)	-4 (Not calculable)	26

TG: Triglycerides; TC: Total cholesterol; HDL: High-density lipoprotein cholesterol; Non-HDL: Total cholesterol minus HDL; Beta-quant LDL-C: LDL cholesterol measured by ultracentrifugation; na: not available

Discussion

These two pediatric cases illustrate the limitations of conventional LDL-C estimation methods in extreme or atypical lipid contexts. Case 1 demonstrates that low total cholesterol combined with moderately elevated triglycerides can produce negative or non-calculable LDL-C values using the Friedewald equation, even when TG levels are well below the classical 400 mg/dL threshold. Alternative formulas, including the Martin–Hopkins and Sampson–NIH equations yielded calculable LDL-C values, but all substantially underestimated LDL-C compared with beta-quantification, with Friedewald showing the most pronounced negative bias. This observation suggests that disproportionately low TC, rather than moderately elevated TG alone, can precipitate calculation failure. Instability occurs because the fixed TG ÷ 5 assumption overestimates VLDL-C when triglycerides are disproportionately high relative to total cholesterol. As the triglyceride-to-total cholesterol (TG/TC) ratio approaches 5, the likelihood of equation failure increases; in Case 1, the TG/TC ratio is 4.6, placing it near this instability threshold (Table 1). From a pathophysiologic standpoint, poorly controlled HIV and AIDS contribute to impaired hepatic lipoprotein synthesis, increased VLDL-TG secretion, and reduced lipoprotein lipase activity, resulting in low LDL-C and moderately elevated TG [10-12]. Together, these mechanisms disrupt the expected relationship between TG and TC, underscoring that the TG–TC balance, rather than absolute TG alone, is central to understanding equation performance and limitations (Table 2). If validated in larger and more diverse pediatric cohorts, the TG/TC ratio may serve as an adjunctive laboratory flag to prompt reflex direct LDL-C measurement, apolipoprotein B testing, or preferential reporting of non-HDL-C.

Case 2 underscores the impact of extreme hypertriglyceridemia on LDL-C estimation. TG levels exceeding 1400 mg/dL, likely driven by asparaginase therapy and insulin resistance, disrupted the TG-TC relationship, rendering the Friedewald and Martin–Hopkins equations non-calculable. In contrast, the Sampson–NIH 2025 formula produced a calculable LDL-C value, demonstrating improved equation stability in extreme lipid perturbations. However, in the absence of beta-quantification, the analytical accuracy of this estimate cannot be determined. A key observation from both cases is that TG/TC ratios exceeding 4.6 (Case 1) and 6.4 (Case 2) were associated with calculation instability in the Friedewald and Martin-Hopkins equations respectively (Table 2). Relying solely on calculated LDL-C in these contexts risks underestimating LDL particle burden and misclassifying lipid profiles, particularly in pediatric and medically complex populations [9,13]. While the modified Sampson–NIH (2025) equation appears more resilient in calculating a LDL value at high TG and low TC setting, caution is still warranted at extreme lipid perturbations. Direct measurement by Beta-quantification, or assessment of apolipoprotein B remains the gold standard for accurate LDL-C evaluation under such extreme physiological conditions. In summary, these cases highlight: (1) Friedewald LDL-C calculation can fail in the setting of low TC even if TG <300 mg/dL producing negative or severely underestimated values. Elevated TG/TC ratios rather than absolute TG alone may be associated with increased likelihood of LDL-C calculation failure. (2) In high TG scenario, alternative equations, including Martin–Hopkins and modified Sampson–NIH (2025) equation, may generate calculable LDL-C value but may not fully eliminate bias, and Martin equation fails to obtain a calculable value at TG/TC>6.4. (3) Direct LDL-C measurement, apolipoprotein B or non-HDL-C markers remain the benchmark in pediatric patients with atypical lipid profiles, systemic illness or complex metabolic states. Clinicians should interpret calculated LDL-C with caution, considering TG/TC ratios and underlying physiologic context.

Table 2: Predictors of LDL-C calculation reliability in extreme lipid contexts.

Situation	Equations reliability to calculate LDL
1. TG ≥ 400 (classic cutoff) 2. Suggested TG/TC > 4.6 (even if TG <400)	Friedewald incalculable Most equations underestimate LDL-C at high TG/TC ratio
1. TG>800 2. Suggested TG/TC > 6.4	Martin incalculable Sampson–NIH 2025 provides calculable LDL-C

Conclusion

Case 1 challenges the conventional view that Friedewald LDL-C calculation fails only in extreme hypertriglyceridemia. Our findings suggest that the TG/TC ratio may represent a potential marker of increased likelihood of calculation failure and requires further validation. Even with TG below 400 mg/dL, low TC combined with moderately high triglycerides (>200 mg/dL) can produce severe negative value with Friedewald calculations, demonstrating that low total cholesterol, rather than hypertriglyceridemia alone, can produce calculation failure. Inthis setting, alternative equations, including Martin–Hopkins and Sampson–NIH 2025, generated calculable values but produced negative bias compared to measured beta-quantification. Clinically, such underestimation may obscure residual atherogenic LDL, including small dense particles, and lead to misclassification of lipid status. At very high TG levels, Martin–Hopkins may also become non-calculable, whereas the Sampson–NIH 2025 equation maintains calculability, although accuracy could not be confirmed, as observed in Case 2. Direct homogeneous LDL-C assays represent an alternative approach;however, recent data suggest potential positive

bias at low LDL-C concentrations when compared with beta-quantification (14). Therefore, in high TG/TC scenarios, the modified Sampson–NIH 2025 equation may provide a practical approach when direct measurement is unavailable, while non-HDL-C remains a robust TG-independent marker of atherogenic burden. Clinicians should therefore interpret low calculated LDL-C with caution in the setting of elevated TG and low TC, especially in patients receiving cholesterol-lowering therapies, where calculated LDL-C values may substantially underestimate true atherogenic burden. Further validation in larger pediatric cohorts with extreme lipid states is ongoing.

Author Statement

Dr. Earnest JP Daniel & Dr. Sridevi Devaraj contributed to the conceptualization and design of the case study, interpretation of results, and provided critical revisions to manuscript. Dr. Sridevi Devaraj, an expert in cardiovascular research is the corresponding author providing expert insight into the LDL related findings and relevant clinical implications.

Funding Statement

This study was not funded by any external sources. The authors declare no financial support for the research, authorship, or publication of this case study. Earnest JP Daniel was supported by an endowment Ching Nan Ou fellowship in Clinical Chemistry at TCH.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon request. The dataset includes clinical records, laboratory results, and other relevant data used in the manuscript

Artificial Intelligence (AI) usage statement

No AI tools were used in the data analysis or writing of this case study. Clinical data were obtained using EPIC Beaker tools. All interpretations and conclusions were made by the authors in accordance with ethical standards, and no automated pattern-finding tools were employed for analysis or case report writing. AI tools were used solely for grammar and language review.

Ethics Statement

This case report was conducted in accordance with institutional ethical guidelines.

Conflict of Interest

The authors declare no conflicts of interest relevant to this work.

References

1. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem. 1972; 18(6):499-502.
2. Martin SS, Blaha MJ, Elshazly MB, Brinton EA, Toth PP, McEvoy JW et al. Friedewald-estimated versus directly measured low-density lipoprotein cholesterol and treatment implications. J Am Coll Cardiol. 2013; 20;62(8):732-739.
3. Sajja A, Park J, Sathiyakumar V, Varghese B, Pallazola VA, Marvel FA, et al. Comparison of Methods to Estimate Low-Density Lipoprotein Cholesterol in Patients With High Triglyceride Levels. JAMA Netw Open. 2021;1;4(10):e2128817.
4. Samuel C, Park J, Sajja A, Michos ED, Blumenthal RS, Jones SR, Martin SS. Accuracy of 23 Equations for Estimating LDL Cholesterol in a Clinical Laboratory Database of 5,051,467 Patients. Glob Heart. 2023;19;18(1):36.
5. White-Al Habeeb N, Agbor T, Sun D, Wan A, Seaton M, Catomeris P, Konforte D. Impact of the implementation of the Sampson-NIH equation for the calculation of LDL-C in a large community cohort of adult and pediatric patients in Ontario. Clin Biochem. 2026; 141:111071.
6. Wang I, Rahman MH, Hou S, Lin HW. Assessing the Practical Differences in LDL-C Estimates Calculated by Friedewald, Martin/Hopkins, or NIH Equation 2: An Observation Cross-Sectional Study. J Lipid Atheroscler. 2023;12(3):252-266
7. Sampson M, Ling C, Sun Q, Harb R, Ashmaig M, Warnick R, et al. A New Equation for Calculation of Low-Density Lipoprotein Cholesterol in Patients With Normolipidemia and/or Hypertriglyceridemia. JAMA Cardiol. 2020;1;5(5):540-548
8. Sampson M, Zubiran R, Wolska A, Meeusen JW, Donato LJ, Jaffe AS, et al. A Modified Sampson-NIH Equation with Improved Accuracy for Estimating Low Levels of Low-Density Lipoprotein-Cholesterol. Clin Chem. 2025;4;71(11):1125-1137.
9. Meeusen JW, Yi X, Cotten SW, Nielsen JB, Donato LJ, Jones PM, Muthukumar AR et al. Modern Low-Density Lipoprotein Cholesterol Formulas Outperform Direct Methods in Patients with Hypertriglyceridemia and Low Levels of Low-Density Lipoprotein Cholesterol. Clin Chem. 2025;4;71(11):1138-1146.
10. Grunfeld C. Dyslipidemia and its Treatment in HIV Infection. Top HIV Med. 2010;18(3):112-118. PMID: 20921577; PMCID: PMC3189481
11. Grunfeld C, Kotler DP, Hamadeh R, Tierney A, Wang J, Pierson RN. Hypertriglyceridemia in the acquired immunodeficiency syndrome. Am J Med. 1989;86(1):27-31. doi: 10.1016/0002-9343(89)90225-8. PMID: 2910092.
12. El-Sadr WM, Mullin CM, Carr A, Gibert C, Rappoport C, Visnegarwala F, Grunfeld C, Raghavan SS. Effects of HIV disease on lipid, glucose and insulin levels: results from a large antiretroviral-naive cohort. HIV Med. 2005;6(2):114-121. doi: 10.1111/j.1468-1293.2005.00273.x. PMID: 15807717.
13. Guirguis-Blake JM, Evans CV, Coppola EL, Redmond N, Perdue LA. Screening for Lipid Disorders in Children and Adolescents: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. 2023;18;330(3):261-274.

14. Gcingca T, Sampson M, Zubirán R, Wolska A, Meeusen J, Donato L, Jaffe A, Remaley A. Lipid Ratio Plot: A Simple Graphical Approach for Investigating the Accuracy of LDL Cholesterol Equations or Direct Assays. J Appl Lab Med. 2025;10(5):1154-1167. doi: 10.1093/jalm/jfaf093. PMID: 40696992.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

Drug-Induced Liver Injury During First-Line Anti-Tubercular Therapy (Isoniazid, Rifampicin, Pyrazinamide, and Ethambutol): A Case Report with De-challenge and Rechallenge

Sridevi Chokkakula[1*], Ashwitha Balasani[1], Sathwika Reddy Ammana[2]

Article Info

Corresponding Author:
Sridevi Chokkakula
Department of Pharmacy Practice, Malla Reddy Pharmacy College, Maisammaguda, Hyderabad, Telangana, India.
E-mail: sridevisamireddy92@gmail.com
Phone: 7780659802

Keywords

ADR (adverse drug reaction), Anti-tubercular therapy, De-challenging, HRZE (isoniazid, Rifampicin, Pyrazinamide, Ethambutol), Re-challenging, TB (Tuberculosis)

Abstract

Background: The primary treatment for tuberculosis, which includes isoniazid, rifampicin, pyrazinamide, and ethambutol, is vital for managing the disease but often leads to drug-induced liver injury (DILI), especially in older adults. It is important to accurately identify and manage hepatotoxicity caused by this treatment to safely continue therapy.

Objective: This report discusses a case of liver injury induced by first-line anti-tubercular drugs in an elderly patient, emphasizing the importance of de-challenge, alternative non-hepatotoxic treatments, and a carefully monitored stepwise rechallenge.

Methods: An 84-year-old woman with pulmonary tuberculosis and hypertension (treated with amlodipine) experienced abnormal liver function tests after starting first-line anti-tubercular therapy. A thorough clinical evaluation, laboratory tests, causality assessment, de-challenge, alternative treatment, and stepwise rechallenge were conducted with close biochemical monitoring.

Results: Initial tests showed significant hyperbilirubinemia (total bilirubin 5.2 mg/dL), elevated aspartate aminotransferase (AST 176–247 U/L), slightly increased Alanine Aminotransferase (ALT 40–55 U/L), and an R-ratio of about 2.6, indicating a predominantly cholestatic pattern with mixed features of liver injury. After discontinuing the hepatotoxic treatment and starting alternative medications (Moxifloxacin, Streptomycin, and Ethambutol), liver function tests improved significantly within 3–5 days. Once liver parameters normalized, Rifampicin was gradually reintroduced, and liver function tests remained stable four days after rechallenge, with no return of hepatotoxicity.

Conclusion: This case illustrates that early detection of liver injury from anti-tubercular therapy, prompt de-challenge, use of alternative non-hepatotoxic drugs, and a carefully monitored stepwise rechallenge can enable the safe continuation of tuberculosis treatment in elderly patients. Close biochemical monitoring, along with consideration of therapeutic drug monitoring, may further improve patient safety.

Introduction

Tuberculosis ranks as a significant infectious disease globally and is one of the top causes of illness and death around the world. The World Health Organization reports that millions of new tuberculosis cases emerge each year, especially in low- and middle-income nations. Successful treatment necessitates extended multidrug therapy to ensure recovery and prevent the development of resistance [11]. "Hepatitis" is a term referring to inflammation of the liver, a condition that can result from various causes, including viral infections, alcohol consumption, drug reactions, or autoimmune disorders. It's a swelling of the liver tissue, which can damage its function [1]. ATT (antitubercular therapy)-induced drug-induced liver injury is defined as an increase in liver enzymes such as alanine transaminase (ALT) and aspartate transaminase (AST) to more than 2-3 times the upper normal limit in individuals taking antitubercular medication in the absence of other causes of liver injury [2]. Among the first-line drugs, isoniazid (H), rifampicin (R), and pyrazinamide (P) have potential for hepatotoxicity, with pyrazinamide being the most hepatotoxic, followed by isoniazid and rifampicin [3]. ATT-induced hepatitis is due to metabolic idiosyncrasy, where the metabolites are released and accumulated during the metabolic process; this may be facilitated by polymorphism of metabolizing enzymes [4]. The pathogenesis of isoniazid-induced hepatotoxicity is not fully understood and involves hepatic metabolism. Isoniazid is metabolized by N-acetyltransferase (NAT2) and CYP2E1 to form acetyl hydrazine and other intermediates. These metabolites produce reactive toxic species that bind to hepatic macromolecules, leading to liver injury [5]. Rifampicin enhances the hepatotoxicity of other anti-tubercular drugs. It activates the pregnane X receptor, leading to induction of drug-metabolizing enzymes such as CYP3A4. This increases isoniazid metabolism and the formation of toxic metabolites, contributing to liver injury. [6]. Rifampicin also induces isoniazid hydrolases, leading to increased hydrazine levels in slow acetylators, thus increasing toxicity [7]. Compared to other primary antitubercular medications, ethambutol-induced liver toxicity is infrequent and unusual. The precise mechanism remains unclear, and ethambutol is not regarded as a direct intrinsic liver toxin. Current evidence indicates that liver damage is primarily idiosyncratic and immune-mediated, rather than related to dosage. Moxifloxacin as an alternative anti-tubercular drug works by inhibiting DNA gyrase (topoisomerase II) and topoisomerase IV, which are crucial for bacterial DNA replication and transcription, thereby preventing the growth of Mycobacterium tuberculosis. Streptomycin is utilized as a second-line tuberculosis medication, primarily targeting extracellular mycobacteria, and is used when hepatotoxic drugs need to be avoided. It irreversibly attaches to the 30S ribosomal subunit, leading to mRNA misreading and blocking protein synthesis, which results in the death of bacterial bactericidal effect [6,8]. Hepatotoxicity is often observed as part of hypersensitivity reactions, such as DRESS syndrome, suggesting an immunological component. Since ethambutol is typically used in combination with other drugs, determining its specific role in liver damage is challenging [12]. Pyrazinamide leads to liver damage primarily due to its dose-dependent metabolic toxicity. Inthe liver, it is converted to pyrazinoic acid and subsequently to 5-hydroxypyrazinoic acid, metabolites that are directly toxic to hepatocytes. These metabolites induce oxidative stress, disrupt mitochondrial function, and impair cellular energy production, resulting in hepatocellular injury and necrosis, which clinically manifests as elevations in alanine aminotransferase and aspartate aminotransferase levels resembling acute hepatitis. Compared to other first-line antitubercular drugs, pyrazinamide is more hepatotoxic because its liver injury is strongly dose dependent, occurs through direct metabolic toxicity rather than idiosyncratic reactions, and lacks a safe metabolic detoxification pathway. The risk is further amplified in elderly patients, those with pre-existing liver disease, and when pyrazinamide is used in combination with other hepatotoxic antitubercular agents [13]. Antitubercular drug–induced liver injury occurs due to multiple biochemical mechanisms. During hepatic metabolism, these drugs form reactive metabolites that can damage mitochondria, leading to reduced energy production and hepatocyte injury. These metabolites induce oxidative stress by generating reactive oxygen species, leading to lipid peroxidation and cell membrane damage. They may also trigger immune-mediated liver injury, with combined effects of mitochondrial dysfunction, oxidative stress, and immune activation contributing to hepatocellular or cholestatic injury [10]. The various risk factors for ATT (anti-tubercular therapy)-induced hepatitis include advanced age, female gender, pregnancy, comorbidities such as diabetes and obesity, underlying liver diseases, genetic factors such as acetylator polymorphism, concomitant viral infections such as HIV and hepatitis B and C, and underlying nutritional status such as malnutrition [8].

While liver injury caused by anti-tubercular drugs is well recognized, there is a scarcity of detailed accounts on how it is managed in older patients, especially concerning the use of alternative medications and the careful, monitored reintroduction of first-line treatments. This case report aims to detail a liver injury caused by medication in an elderly patient undergoing first-line anti-tubercular treatment. It also emphasizes the importance of dechallenge, the use of alternative therapies that do not harm the liver, and a carefully monitored step-by-step rechallenge to ensure the safe continuation of tuberculosis treatment.

Materials and methods

Study Design

This manuscript is a descriptive, single-patient case report documenting drug-induced liver injury associated with first-line anti-tubercular therapy.

Study Setting and Duration

The case was identified and managed in the Department of General Medicine at Malla Reddy Hospital, Suraram, Hyderabad, Telangana, India. The clinical evaluation, management, and follow-up were carried out during the patient's hospital admission in 2024.

Patient Selection

The patient was selected based on the clinical presentation of

abnormal liver function tests following initiation of first-line anti-tubercular therapy.

Inclusion Criteria

Patients were included if they:
We're receiving first-line anti-tubercular therapy (isoniazid, rifampicin, pyrazinamide, and ethambutol)
Developed abnormal liver function tests temporally associated with anti-tubercular drug exposure
Showed improvement in liver function tests following dechallenge of suspected hepatotoxic drugs

Exclusion Criteria

Patients were excluded if they:
Had pre-existing chronic liver disease
Had confirmed viral hepatitis or obstructive hepatobiliary disease as the primary cause of liver dysfunction
Had liver injury attributable to causes other than anti-tubercular drugs

Case

An 84-year-old woman was admitted to the General Medicine department with complaints of loose stools, vomiting, poor oral intake, and epigastric abdominal pain for the past three days. She had been diagnosed with pulmonary tuberculosis two months earlier and was receiving first-line anti-tubercular therapy comprising isoniazid, rifampicin, pyrazinamide, and ethambutol. She was also a known hypertensive for several years and was on regular treatment with amlodipine 5 mg once daily. The therapy was temporarily discontinued during the evaluation of liver injuryas a precautionary measure in the context of polypharmacy, although it is not commonly associated with hepatotoxicity. There was no history of chronic liver disease, alcohol use, or intake of other hepatotoxic medications.

Lab investigations

Her lab data showed IgG Ab positive (8.81). A normal value of <1.00 to rule out infectious causes of liver damage, viral hepatitis screening was conducted. The test for Hepatitis A IgM came back negative, while Hepatitis A IgG was positive (8.81), suggesting previous exposure and immunity rather than a current infection. Tests for Hepatitis B surface antigen (HBsAg) and anti-Hepatitis C virus (anti-HCV) antibodies were also negative. These results effectively exclude acute viral hepatitis as the source of the liver injury. Due to the swift biochemical recovery after stopping the medication and the clear timing with antitubercular treatment, autoimmune markers were not assessed, as the rapid biochemical improvement following drug withdrawal strongly supported a drug-induced etiology, making alternative causes such as autoimmune hepatitis less likely. An abdominal ultrasound revealed that the liver's echotexture was normal, with no signs of biliary blockage, gallstones, or dilation within the intrahepatic bile ducts; this excludes obstructive cholestasis. Reduced Hb levels to 9.7 gm% and reduced RBC levels to 3.1 m/cmm with an increased ESR. Liver function tests showing the rise in ALT, AST, and bilirubin levels and decreased total protein levels. Electrolytes showing hyponatremia and hypokalemia with urine culture positive for E. coli. Child-Pugh test score 5. Renal function tests were found to be normal.

Table 1: Liver function test results at baseline before de-challenge of first-line anti-tubercular therapy.

LFT	Day1	Day2	Reference Values
Total Bilirubin	5.2	5.0	0.3-1.3mg/dL
Direct Bilirubin	3.1	3.8	0.1-0.3mg/dL
Indirect Bilirubin	2.1	1.2	0.2-0.8md/dL
AST	176	247	10-40U/L
ALT	40	55	7-56U/L
ALP	63	63	44-147U/L
Total protein	5.6	5.4	6-8.3g/dL
Albumin	3.2	3.1	3.5-6.5g/dL
Globulin	2.4	2.3	2-3.5g/dL
A/G	1.33	1.37	1.5-2.5

Table 2: Liver function test trends following de-challenge and treatment with alternative anti-tubercular agents (Day 3–Day 5).

LFT	Day3	Day4	Day5	Reference Values
Total Bilirubin	1.8	1.8	1.6	0.3-1.3mg/dL
Direct Bilirubin	1.2	1.2	1.1	0.1-0.3mg/dL
Indirect Bilirubin	0.6	0.6	0.5	0.2-0.8md/dL

LFT	Day3	Day4	Day5	Reference Values
AST	89	59	33	10-40U/L
ALT	68	61	38	7-56U/L
ALP	57	60	53	44-147U/L
Total protein	5.3	5.7	5.4	6-8.3g/dL
Albumin	3.0	3.3	3.2	3.5-6.5g/dL
Globulin	2.3	2.4	2.3	2-3.5g/dL
A/G	1.30	1.37	1.39	1.5-2.5

Table 3: Liver function test results four days after stepwise rechallenge with rifampicin.

LFT	Four days after administration of rifampicin	Reference Values
Total Bilirubin	1.6	0.3-1.3mg/dL
Direct Bilirubin	1.4	0.1-0.3mg/dL
Indirect Bilirubin	0.2	0.2-0.8md/dL
AST	24	10-40U/L
ALT	17	7-56U/L
ALP	54	44-147U/L
Total protein	5.6	6-8.3g/dL
Albumin	3.2	3.5-6.5g/dL
Globulin	2.4	2-3.5g/dL
A/G	1.33	1.5-2.5

Table 4: Showing Serum Electrolyte Levels.

Electrolytes	DAY 1	DAY 4	Reference Values
Sodium	131 mEq / L	136 mEq/L	135-145mEq/L
Potassium	2.5 mEq/L	3.6 mEq/L	3.5-5mEq/L
Chloride	97 mEq/L	97 mEq/L	98-106mEq/L

Upon initial examination, liver function tests revealed marked hyperbilirubinemia, with a total bilirubin concentration of 5.2 mg/dL, alongside increased aspartate aminotransferase (AST) and slightly raised alanine aminotransferase (ALT) levels. To characterize the pattern of liver injury, the R-ratio was calculated using the formula R = (ALT/ULN) ÷ (ALP/ULN). The peak ALT level was 55 U/L (ULN 40 U/L), and ALP was 63 U/L (ULN 120 U/L). The calculated R-ratio was therefore (55/40) ÷ (63/120) ≈ 2.63, indicating a mixed hepatocellular–cholestatic pattern of liver injury. This biochemical pattern, along with marked hyperbilirubinemia and modest transaminase elevation, supported a predominantly cholestatic pattern with mixed features,consistent with drug-induced liver injury. This was corroborated by the significant rise in bilirubin levels, modest elevation in transaminases, and nearly normal ALP levels. After discontinuing the suspected hepatotoxic antitubercular medications, there was a gradual return to normal of bilirubin and liver enzyme levels, indicating a reversible drug-induced liver injury. Due to abnormal liver function tests indicating potential liver damage from anti-tubercular drugs, all first-line anti-tubercular medications (isoniazid, rifampicin, and pyrazinamide) with hepatotoxic potential were discontinued (dechallenge). The patient was then started on alternative anti-tubercular drugs that are not harmful to the liver, such as moxifloxacin, streptomycin, and ethambutol.

Liver function was closely monitored, and within three to five days of beginning the new treatment, there was a noticeable improvement in serum bilirubin and transaminase levels, suggesting a successful de-challenge response. After liver function tests returned to normal, the initial anti-tubercular treatment was gradually reintroduced, beginning with rifampicin, and no signs of hepatotoxicity reappeared during the first fourdays of this process.

Initial serum electrolyte tests showed slight hyponatremia (Na: 133 mEq/L) and hypokalemia (K: 2.5 mEq/L), which improved upon retesting (Na: 134 mEq/L, K: 3.6 mEq/L) after supportive care. These imbalances were linked to gastrointestinal losses and decreased oral intake and were resolved before resuming anti-tubercular treatment.

Table 5: RUCAM Causality Assessment for Anti-Tubercular Therapy-Induced Liver Injury.

RUCAM Criteria	Clinical Findings in present case	Score
Time of onset of liver injury after starting the drug	Patient developed abnormal liver function tests approximately 2 months after initiation of ATT, which falls within the typical time frame for ATT-Induced DILI	+2
Course of ALT after stopping the drug (DECHALLENGE)	ALT and Bilirubin levels improved significantly within 3-5 days after withdrawal of hepatotoxic drugs	+3
Risk factors	Age>55 years (84 years)	+1
Concomitant drugs	Amlodipine was used but discontinued and not strongly associated with significant hepatotoxicity	0
Exclusion of non-drug causes	Viral hepatitis screening performed (Hepatitis A IgM negative, HBsAg negative, Anti-HCV negative); no history of alcohol use or chronic liver disease; imaging excluded biliary obstruction.	+2
Previous hepatotoxicity information for the drug	Hepatotoxicity of Isoniazid, Rifampicin, and Pyrazinamide is well documented in literature	+2
Response to re-administration (RECHALLENGE)	Rifampicin rechallenge did not reproduce hepatotoxicity during the monitored period.	0
TOTAL SCORE		10

Causality assessment was performed using the RUCAM (Roussel Uclaf Causality Assessment Method) scale, which yielded a score of 10, indicating highly probable drug-induced liver injury.

Treatment

In view of markedly abnormal liver function tests suggestive of drug-induced liver injury, first-line anti-tubercular therapy was withheld (dechallenge). Alternative treatment with non-hepatotoxic drugs of Inj. MOXIFLOXACIN 400 mg IV OD, initiated as an alternative treatment for tuberculosis, particularly when primary medications are unsuitable due to side effects or drug resistance. Inj. STREPTOMYCIN 0.75 g IM OD is utilized as a second-line tuberculosis medication, primarily targeting extracellular mycobacteria, and is used when hepatotoxic drugs need to be avoided. and Inj. ETHAMBUTOL 800 mg is employed as a supplementary antitubercular drug to avert drug resistance and boost treatment effectiveness. It is given along with liver protectants such as Inj.

L-ORNITHINE-L-ASPERTATE 6 amp in 10 DNS slow IV over 8 hours, T. URSODEOXYCHOLIC ACID 300 mg PO BD, and T. GLUTATHIONE 500 mg PO BD. On the third day the patient was administered Inj. N-ACETYL CYSTINE 600 mg PO BD.

On the fifth day, MOXIFLOXACIN was stopped, and T. LEVOFLOXACIN 500 mg OD was administered. Liver function tests were performed until optimal results were obtained.

After normalization of liver function tests, a cautious stepwise rechallenge of first-line anti-tubercular therapy was initiated, beginning with low-dose rifampicin under close biochemical monitoring: RIFAMPICIN 150 mg on day 1, and the dose was incrementally increased to T. RIFAMPICIN 300 mg on day 2 and 400 mg on day 4. Along with FAROPENAM 200 mg P/O OD, T. LEVOFLOXACIN and Inj. STREPTOMYCIN were withheld, and LFTs were repeated, and they showed normal results. The ophthalmologist's opinion was taken, and a fundoscopy was performed: There is no evidence was found of retinopathy changes; a psychiatrist's opinion was taken in view of the past history of delirium. They advised T. ZOLPIDEM 5 mg SOS (when the patient is unable to sleep).

DISCHARGE MEDICATION

T. RIFAMPICIN 450mg PO OD

T. ETHAMBUTOL 800mg PO OD

T. URSODEOXYCHOLIC ACID 300mg PO BD x 5 days

T.PANTOPRAZOLE 40mg PO OD x 5 days
T. ZOLPIDEM 5mg PO SOS

Discussion

Drug-induced liver injury is a known complication of first-line antitubercular therapy, especially in older patients who may experience altered drug metabolism and diminished liver function. In this case, the liver injury's biochemical pattern, its timing with the start of antitubercular therapy, and the quick recovery after stopping the medication strongly indicate ATT-associated DILI. As the patient also had a urinary tract infection caused by Escherichia coli and electrolyte disturbances, including hyponatremia and hypokalemia, likely related to gastrointestinal losses and reduced oral intake. Although systemic infection and metabolic disturbances can occasionally influence liver enzyme levels, the rapid normalization of liver function tests following withdrawal of hepatotoxic drugs suggests that antitubercular therapy remained the most plausible cause of liver injury in this case. The successful gradual reintroduction of rifampicin without the return of liver toxicity further confirmed the diagnosis while enabling the continuation of tuberculosis treatment.

The treatment for ATT-induced liver diseases should be initiated at an appropriate time; if not, this may lead to life-threatening conditions such as resistant TB. Once the diagnosis of ATT-induced hepatitis is established, it is essential to first stop all potentially hepatotoxic drugs till the LFTs become normal. In the interim period, at least three non-hepatotoxic drugs, such as ethambutol, streptomycin, and quinolones such as levofloxacin, ofloxacin, and moxifloxacin, can be used after evaluation of renal function and visual acuity.

The liver injury pattern observed in this patient is consistent with hepatotoxicity associated with first-line antitubercular therapy. The marked elevation in bilirubin levels with relatively lower transaminases suggests a predominantly cholestatic pattern rather than hepatocellular injury. The patient's clinical improvement following drug withdrawal, along with the absence of worsening upon cautious rechallenge, further supports the diagnosis of drug-induced liver injury. These findings highlight the importance of early recognition and timely management to prevent complications.

According to National Tuberculosis Eradication Programme (NTEP) guidelines and recommendations from the World Health Organization (WHO), rechallenge with antitubercular drugs is generally initiated once transaminase levels return to less than two times the upper limit of normal, with stepwise reintroduction of drugs and adequate monitoring intervals. In the present case, a rechallenge strategy was implemented; however, the duration of monitoring between drug reintroductions was shorter than typically recommended in standard guidelines. This deviation was a pragmatic clinical decision, necessitated by the patient's advanced age and the urgent need to resume effective tuberculosis therapy. Rifampicin was selected as the initial drug for rechallenge due to its comparatively lower hepatotoxic potential relative to isoniazid and pyrazinamide and its established role as the backbone of antitubercular therapy, in alignment with national and WHO guidance. However, rechallenge was limited to rifampicin alone, and isoniazid and pyrazinamide were not reintroduced. This limits definitiveattribution of hepatotoxicity to a specific agent. Despite the shortened rechallenge period, close clinical and biochemical monitoring was performed to ensure patient safety. The approach reflects a balance between guideline recommendations and real-world clinical urgency. Based on the WHO-UMC causality assessment scale, this case was categorized as a probable adverse drug reaction.

The World Health Organization–Uppsala Monitoring Centre (WHO-UMC) causality assessment scale was utilized to evaluate causality. A distinct temporal link was identified between the start of first-line antitubercular treatment and the emergence of liver dysfunction. Clinical and laboratory assessments ruled out other potential causes of liver damage, such as viral hepatitis and obstructive hepatobiliary disease. During a stepwise rechallenge, rifampicin was reintroduced, and liver function tests remained normal four days after reintroduction, with no recurrence of hepatotoxicity. Given the temporal correlation, positive dechallenge outcome, and exclusion of other causes, the reaction was deemed a probable adverse drug reaction according to WHO-UMC criteria. A structured causality assessment using the Roussel Uclaf Causality Assessment Method (RUCAM) yielded a score, indicating a highly probable drug-induced liver injury. Most studies published on liver injury caused by anti-tubercular drugs focus on mixed-age or relatively younger groups and mainly report hepatocellular or mixed liver injury patterns. These studies typically involve stopping the problematic drugs, with limited documentation of stepwise rechallenge, as it is often avoided due to safety concerns. In contrast, the current case involves a very elderly patient (84 years old), a demographic often under-represented in the literature, and shows a mixed liver injury pattern was observed, with features suggesting a predominantly cholestatic component, while most reports focus on hepatocellular injury. Additionally, this case includes a well-documented laboratory timeline, showing baseline abnormalities, improvement after de-challenge from Day 3 to Day 5, and stable liver function tests four days after a stepwise rechallenge. The successful continuation of first-line anti-tubercular therapy following a carefully monitored rechallenge underscores a practical management strategy that goes beyond diagnosis and may provide valuable clinical guidance.

With the stated objective, this case highlights that cautious rifampicin rechallenge may be feasible even in very elderly patients with suspected anti-tubercular drug-induced liver injury when careful monitoring is undertaken.

Limitations

This report is constrained by its focus on a single patient, which limits its applicability to a broader population. Additionally, procedures such as liver biopsy, therapeutic drug monitoring, and pharmacogenetic testing (such as NAT2 polymorphism analysis) were not conducted. Furthermore, the multiple hepatotoxic antitubercular drugs were discontinued simultaneously, making it difficult to identify the causative agent, and the rechallenge period was relatively short. Despite these constraints, the case offers a clear temporal relationship, and a positive dechallenge

response, exclusion of alternative causes, and supportivecausality assessment provide reasonable evidence and also provide valuable clinical insights into managing anti-tubercular drug-induced liver injury in an elderly patient.

Conclusion

This case report discusses a drug-induced liver injury linked to first-line anti-tubercular treatment in an elderly patient, characterized by a cholestatic liver injury pattern. Early identification, swift withdrawal of hepatotoxic medications, and the introduction of alternative non-hepatotoxic anti-tubercular drugs led to a quick biochemical recovery. A cautious, gradual reintroduction of rifampicin with close laboratory monitoring was well-tolerated without a return of hepatotoxicity, allowing for the safe continuation of tuberculosis therapy. This case emphasizes the significance of personalized management strategies for elderly patients and highlights the importance of close biochemical monitoring and therapeutic drug monitoring as essential tools to improve the safety of de-challenge–rechallenge strategies in cases of anti-tubercular drug-induced liver injury.

Conflict of interest

None.

Funding Statement

Nil.

Ethical Approval Statement

This paper presents a case study involving a single patient, based on standard clinical practice. According to the ethical guidelines of Malla Reddy Hospital in Suraram, Hyderabad, Telangana, India, individual case reports are considered clinical documentation rather than research involving human subjects and thus do not require a formal reference number from the Institutional Ethics Committee. The committee has reviewed and approved the publication of this case. The patient provided written informed consent for the publication.

Patient Consent Statement

Written informed consent was obtained from the patient for publication of the case report.

Acknowledgement

The authors would like to thank the guide, Dr. CH. Sridevi, and Dr. G. Tulja Rani, principal of Malla Reddy Pharmacy College, for their kind support and encouragement, and the physicians of the general medicine department, Malla Reddy Hospital, Suraram.

Credit Author Statement

Author 1*: Supervision, final approval of manuscript, and support in case interpretation.
Author 1: Conceptualization, Data Collection, Clinical Examination, Investigation, Writing-Original Draft Preparation.
Author 2: Literature search, editing and validation of clinical details, reference management, and proofreading.

Data Availability Statement

The data supporting this case report are not publicly available due to patient privacy and confidentiality. Anonymized information may be provided by the corresponding author upon reasonable request.

References

1. Zhao H, Wang Y, Zhang T, Wang Q, Xie W. Drug-induced liver injury from anti-tuberculosis treatment: A retrospective cohort study. Med Sci Monit [Internet]. 2020;26:e920350. Available from: http://dx.doi.org/10.12659/MSM.920350
2. Giannini EG, Testa R, Savarino V. Liver enzyme alteration: a guide for clinicians. CMAJ [Internet]. 2005;172(3):367–79. Available from: http://dx.doi.org/10.1503/cmaj.1040752
3. Hoofnagle JH, Björnsson ES. Drug-induced liver injury - types and phenotypes. N Engl J Med [Internet]. 2019;381(3):264–273. Available from: http://dx.doi.org/10.1056/NEJMra1816149
4. Knowles SR, Uetrecht J, Shear NH. Idiosyncratic drug reactions: the reactive metabolite syndromes. Lancet [Internet]. 2000;356(9241):1587–1591. Available from: http://dx.doi.org/10.1016/S0140-6736(00)03137-8
5. Huang Y-S, Chern H-D, Su W-J, Wu J-C, Lai S-L, Yang S-Y, et al. Polymorphism of the N-acetyltransferase 2 gene as a susceptibility risk factor for antituberculosis drug-induced hepatitis. Hepatology [Internet]. 2002;35(4):883–889. Available from: http://dx.doi.org/10.1053/jhep.2002.32102
6. Shen C, Meng Q, Zhang G, Hu W. Rifampicin exacerbates isoniazid-induced toxicity in human but not in rat hepatocytes in tissue-like cultures. Br J Pharmacol [Internet]. 2008;153(4):784–791. Available from: http://dx.doi.org/10.1038/sj.bjp.0707611
7. Chen J, Raymond K. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. Ann Clin Microbiol Antimicrob [Internet]. 2006;5(1):3. Available from: http://dx.doi.org/10.1186/1476-0711-5-3
8. Anand AC, Seth AK, Paul M, Puri P. Risk factors of hepatotoxicity during anti-tuberculosis treatment. Med J Armed Forces India [Internet]. 2006;62(1):45–49. Available from: http://dx.doi.org/10.1016/S0377-1237(06)80155-3
9. Centers for Disease Control and Prevention. Clinical guidelines for TB. Available from: https://www.cdc.gov/tb/. Accessed: 04/12/2025.
10. Knowles SR, Uetrecht J, Shear NH. Idiosyncratic drug reactions: the reactive metabolite syndromes. Lancet [Internet]. 2000;356(9241):1587–1591. Available from: http://dx.doi.org/10.1016/S0140-6736(00)03137-8
11. Global tuberculosis reports [Internet]. Who.int. [cited 2026 Jan 6]. Available from: https://www.who.int/teams/global-tuberculosis-programme/tb-reports
12. Ethambutol. In: LiverTox: Clinical and Research Information on Drug-Induced Liver Injury. Bethesda (MD): National Institute of Diabetes and Digestive and Kidney Diseases; 2012.

13. Hussain Z, Zhu J, Ma X. Metabolism and hepatotoxicity of pyrazinamide, an antituberculosis drug. Drug Metab Dispos [Internet]. 2021;49(8):679–682. Available from: http://dx.doi.org/10.1124/dmd.121.000389

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

The Hypoxia Trap: Worsening Sickle Cell Disease in the Setting of Tetralogy of Fallot

Mohnish Darshan[1*], Sonal Rajmane[1], Jignesh Sharma[1], Amber Kumar[1], Girish Chandra Bhatt[1], Shikha Malik[1]

[1]Department of Pediatrics, All India Institute of Medical Sciences (AIIMS), Bhopal, Madhya Pradesh, India

Article Info

Corresponding Author:
Mohnish Darshan
Department of Pediatrics, All India Institute of Medical Sciences (AIIMS)
Bhopal, Madhya Pradesh, India, Pin – 462020
E-mail: biomd27@gmail.com
Phone: 9840351074

Keywords

Tetralogy of Fallot, Sickle Cell Disease, Pediatric Stroke, Thrombosis, Exchange Transfusion, Cyanotic Congenital Heart Disease

Abstract

Background: Tetralogy of Fallot (TOF) is the most common cyanotic congenital heart disease and predisposes affected children to thrombotic complications due to chronic hypoxia and polycythemia. Sickle cell disease (SCD) further increases the risk of vaso-occlusion through erythrocyte sickling and endothelial dysfunction. The coexistence of TOF and SCD is rare and creates complex diagnostic and therapeutic challenges with significant risk of catastrophic neurological events.

Case presentation: We report a 14-month-old female with known TOF who presented with acute left-sided hemiparesis, fever, and seizures. Brain computed tomography with angiography revealed extensive bilateral ischemic infarctions with complete thrombosis of both middle cerebral arteries and occlusion of the right internal carotid artery. Haematological evaluation confirmed previously undiagnosed SCD with an HbS fraction of 60.2%. Despite intensive therapy including oxygen, intravenous hydration, hydroxyurea, antiepileptic drugs, and exchange transfusion that reduced HbS to 28.9%, the child developed progressive encephalopathy and died. Chronic hypoxia from TOF combined with SCD-related sickling likely precipitated catastrophic cerebrovascular thrombosis.

Conclusion: This case highlights the synergistic thrombotic risks of TOF and SCD, emphasising the critical need for routine hemoglobinopathy screening in children with cyanotic heart disease, especially in high-prevalence regions. Early diagnosis, proactive use of vasopressors to optimise oxygenation, and multidisciplinary management are essential to improve outcomes. Aggressive preventive interventions are essential to reduce morbidity and mortality in patients presenting with these rare coexisting pathologies. Enhanced screening awareness, stroke-preventing strategies and timely surgical intervention for TOF could improve prognosis in this rare, high-risk dual pathology.

Introduction

Tetralogy of Fallot is the most frequent cyanotic congenital heart disease encountered in pediatric practice and comprises four cardinal anatomical abnormalities: ventricular septal defect, overriding aorta, right ventricular hypertrophy, and right ventricular outflow tract obstruction [1]. Chronic cyanosis, recurrent hypoxic spells, and compensatory polycythemia predispose affected children to thromboembolic complications, including cerebrovascular accidents [1].

Sickle cell disease is a common hereditary hemoglobinopathy in geographically distributed regions such as India and sub-Saharan Africa and is characterized by recurrent erythrocyte sickling, hemolysis, and microvascular occlusion [8,11]. Cerebrovascular disease is a major cause of morbidity and mortality in children with SCD, arising from hypoxia-driven endothelial dysfunction and thrombosis [8,11].

The coexistence of TOF and SCD is exceptionally rare but poses a unique clinical challenge due to synergistic amplification of hypoxia-induced sickling and hyperviscosity-related thrombosis [2,3]. Only a limited number of reports describe this dual pathology, and standardized management guidelines are lacking [3,14].

We report a 14-month-old child with TOF who presented with acute massive bilateral cerebral infarction and was subsequently diagnosed with previously unrecognized SCD. This case underscores the importance of early identification of coexisting thrombotic risk factors, routine hemoglobinopathy screening in cyanotic congenital heart disease, and coordinated multidisciplinary management, particularly in resource-limited settings [3,14].

Case description

A 14-month-old female child was referred to a Tertiary care center in Central India with complaints of acute-onset, non-progressive left-sided hemiparesis for three days. This neurological deficit was associated by high-grade fever, non-projectile vomiting, and an episode of generalized tonic-clonic seizures two days before admission. There were no antecedent trauma, urinary retention, or other complaints suggestive of cranial nerve involvement. Her history was notable for recurrent respiratory tract infections and bluish discolouration of the body and extremities, leading to a diagnosis of Tetralogy of Fallot (TOF) with an associated atrial septal defect (ASD) at theage of 6 months. The child had been started on prophylactic propranolol (1 mg/kg/day), and surgical intervention was planned. Birth and developmental histories were unremarkable, and family history revealed one spontaneous abortion without further significance.

Upon reviewing investigations from the previous hospital, cranial computed tomography (CT) revealed confluent hypodense areas with loss of grey-white differentiation in the right frontoparietal-temporal region, along with a midline shift of 6 mm to the left, which is overall suggestive of acute infarction. Echocardiography demonstrated a significant subaortic ventricular septal defect (VSD) with 50% override, bidirectional shunt, infundibular pulmonary stenosis, and an ostium secundum-type ASD measuring 6 mm.

On admission, the child was hemodynamically stable, with a regular heart rate of 115 beats per minute, a respiratory rate of 35 breaths per minute, blood pressure of 100/60 mmhg (within the 90th–95th percentile for age), and an oxygen saturation of 77% on room air. Physical examination revealed pallor, cyanosis, and a grade 4 systolic ejection murmur at the left parasternal border. Neurological assessment showed a Glasgow Coma Scale (GCS) score of 12/15 (E4V3M5), normal cranial nerve function, left-sided hypotonia, grade 2/5 motor power in the left upper and lower limbs, areflexia, and absent left plantar response. Sensory examination indicated loss of pain, touch, and temperature sensation on the left side. No meningeal signs were present. Abdominal examination showed the presence of firm, non-tender splenomegaly (4 cm) along with hepatomegaly (3 cm).

The patient was admitted to the pediatric intensive care unit and initiated on intravenous fluids (100 ml/kg/day normal saline), supplemental oxygen (3 L/min via nasal cannula), decongestive therapy (hypertonic saline), and anti-seizure medication (levetiracetam 20 mg/kg loading dose, followed by 10 mg/kg twice daily). Given the early presentation of stroke in a patient with TOF, an evaluation for thrombophilia was initiated. Complete blood count showed hemoglobin of 12.5 g/dl, haematocrit of 38%, and reticulocyte count of 6%. Peripheral smear revealed sickled erythrocytes, confirmed by a positive sickling test. Biochemical parameters, coagulation profile, and homocysteine levels were normal. High-performance liquid chromatography (HPLC) confirmed the sickle cell disease, with an HbS fraction of 60.2% (Table 1).

Table 1: Baseline clinical and laboratory investigations at presentation.

Parameter	Baseline Result
HbS	60.2%
HbF	24.6%
HbA0	4.1%
HbA2	1.0%
Sickling test	Positive
Hemoglobin	11.1 g/dl
RBC count	4.46 million/µl

Parameter	Baseline Result
MCV	83.4 fl
MCH	24.9 pg
RDW	31.8%
WBC count	17.0 $\times 10^3/\mu l$
Platelet count	163 $\times 10^3/\mu l$
Reticulocyte count	6.7%
Peripheral Blood Smear	Sickle cells, fragmented RBCs, tear-drop cells, nRBCs
CRP	27.1 mg/L
Sodium	152.9 mmol/L
Potassium	5.41 mmol/L
Creatinine	0.48 mg/dl
Homocysteine	9.16 µmol/L

Initial hematological and biochemical parameters at admission demonstrating elevated HbS fraction, positive sickling test, reticulocytosis, and inflammatory markers, with peripheral smear showing sickled and fragmented erythrocytes consistent with sickle cell disease.

Parental HPLC screening revealed sickle cell trait in both parents. Hydroxyurea (15 mg/kg/day) was started to reduce sickling. Given the high HbS fraction and for secondary stroke prevention, Partial exchange transfusion was performed (around 300ml) under strict asepsis protocol. Post-transfusion HPLC showed an HbS fraction of 28.9% (Table 2).

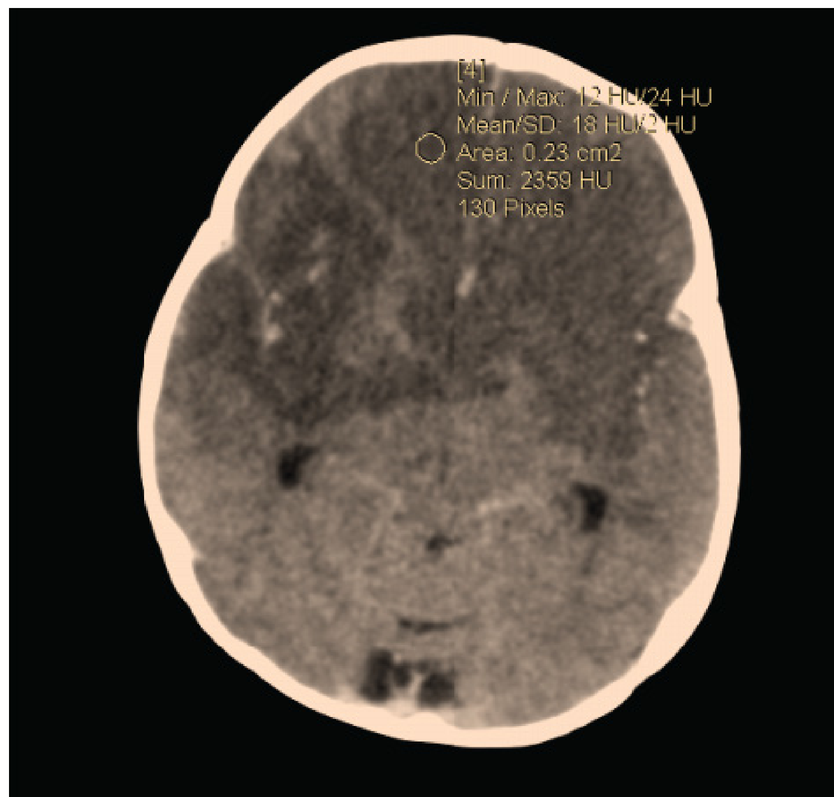
Table 2: Changes in hemoglobin fractions on high-performance liquid chromatography before and after exchange transfusion.

Hemoglobin Fraction	Pre-exchange HPLC (%)	Post-exchange HPLC (%)
Hbs	60.2	28.9
Hbf	24.6	12.8
Hba0	4.1	46.6
Hba2	1.0	1.9

Comparison of hemoglobin fractions demonstrating a marked reduction in HbS from 60.2% to 28.9% following exchange transfusion, with a corresponding increase in HbA0, indicating effective replacement of sickled erythrocytes and improved hemoglobin profile.

However, within hours, the patient developed intractable seizures, requiring escalation of anti-seizure therapy (midazolam infusion 0.1 mg/kg/hour). She progressed to global hypotonia and encephalopathy (GCS 7/15). A Repeat CT head and carotid angiography were performed, which demonstrated a large, relatively defined wedge-shaped hypodense area with loss of grey-white matter differentiation in the bilateral MCA and ACA territory regions, involving the bilateral frontoparietal lobes, basalganglia, and anterior limbs of the internal capsules. Hyperdense bilateral MCA vessels and bilateral A2 segments of the ACA were visualized, consistent with acute-subacute infarction. There was partial effacement of the left lateral ventricle due to mass effect. The angiography revealed complete long-segment non-visualization of the right internal carotid artery (ICA) from its origin, with the absence of the right middle cerebral artery (MCA) and only faint opacification of the A1 segment of the right anterior cerebral artery (ACA), indicating thrombosis. There was also complete non-opacification of the left MCA from its origin, consistent with thrombosis. Other supra-aortic branches, vertebral and basilar arteries, were unremarkable, and the venous sinuses appeared normal in calibre and opacification (Figure 1).

Figure 1: Neuroimaging demonstrating extensive bilateral large-vessel ischemic stroke.



Axial contrast-enhanced computed tomography (CECT) of the brain shows bilateral frontoparietal hypodense infarcts with loss of gray–white matter differentiation involving the middle cerebral artery (MCA) and anterior cerebral artery (ACA) territories. CT angiography reveals long-segment non-visualisation of the right internal carotid artery and bilateral MCAs, consistent with acute large-vessel thrombotic ischemic stroke.

Despite aggressive supportive care, including mechanical ventilation and inotropic support, the patient succumbed to her illness on day 5 of admission. Resource constraints, including limited access to real-time neurological monitoring, bedside palliative shunt expertise, likely contributed to the challenges in managing post-transfusion complications and delayed TOF correction. This case highlights the complex interplay and synergistic and compounding effects of hypoxia due to dual TOF and SCD.

Informed consent for the publication of this case report was obtained from the patient's parents.

Discussion

The unusual co-occurrence of TOF and SCD in a 14-month-old child poses a clinical challenge due to the synergistic thrombotic risks of these conditions [2,3]. TOF results in chronic hypoxia (spo2 77% in our case) and secondary polycythemia (hematocrit 38%), ultimately increasing blood viscosity and predisposing to thrombosis [1,3].

SCD increases this risk through hypoxia-induced erythrocyte sickling, hemolysis, and microvascular occlusion, as seen in the patient's high HbS (60.2%) [8,11]. The intricate interaction of these disorders produced a vicious cycle of hypoxia-inducedsickling, which led to a devastating neurological injury that included bilateral middle cerebral artery thrombosis and a large-scale cerebral infarction [9,12].

The delayed diagnosis of SCD, despite a known TOF diagnosis at 6 months, highlights an important gap in hemoglobinopathy screening practices [3]. In a high-prevalence country like India, routine hemoglobinopathy screening at CCHD diagnosis is essential to detect SCD early [11,14]. Had SCD been detected earlier in this case, preventive interventions such as hydroxyurea, folate supplementation, or pneumococcal vaccination could have reduced the risk of sickling and vaso-occlusive crisis [10,13]. Although the patient's haematocrit is elevated, it indicates that polycythaemia exacerbated hyperviscosity, which was not measured during previous treatment [1]. In individuals with TOF, routine haematocrit and oxygen saturation monitoring may help direct prompt measures to reduce thrombotic risk [4]. The proactive use of vasopressors to maximise oxygenation and minimise sickling is a novel therapeutic approach that has been suggested in this instance. When cyanotic episodes occur in TOF, vasopressors such as phenylephrine are used to raise systemic vascular resistance (SVR), which lessens right-to-left shunt and increases pulmonary blood flow [3,4]. Chronic hypoxia probably caused sickling in this patient, which led to cerebral thrombosis [9]. Proactive vasopressor therapy may have improved systemic oxygenation, reduced sickling, and promoted left-to-right shunting through the VSD by increasing SVR, even outside of acute episodes [5]. Although there is evidence to support the use of vasopressors in the acute management of TOF [3,4], their wider application in TOF-SCD patients to avoid hypoxia-driven consequences has not been investigated. Risks

including elevatedcardiac afterload need to be considered, and pilot studies are required to determine treatment effectiveness [5]. Exchange transfusion helped lower the post-surgical HbS to 28.9%, which is consistent with recommendations for preventing subsequent stroke [6]. However, post-procedural seizures suggest complications, such as hemodynamic shifts or inadequate neurological monitoring, because of the limited availability of real-time EEG [12]. Standardised protocols for exchange transfusion in pediatric SCD-CCHD patients, including accurate volume calculations, continuous monitoring, and pre-procedural stabilisation, could reduce such risks [14]. Limited availability or a lack of patient awareness may be the cause of delays in TOF surgical correction, underscoring the need for robust awareness campaigns and effective capacity building [7]. An earlier approach, such as BT shunt placement or even complete repair, may have improved outcomes [7].

Review of the limited available literature on the coexistence of tetralogy of Fallot and sickle cell disease demonstrates a consistently high burden of morbidity related to thrombotic and neurological complications. Iannucci et al. reported long-term outcomes in five children with sickle cell disease and cyanotic congenital heart disease, including tetralogy of Fallot [2]. They observed frequent Vaso-occlusive events, cerebrovascular complications, and the need for intensive multidisciplinary care. Unlike the patients described in this series, the undiagnosed sickle cell disease in our patient, together with delayed surgical correction of tetralogy of Fallot, likely contributed to the markedly adverse outcome, reinforcing the importance of early evaluation and timely surgical intervention in such high-risk children [7].

To improve outcomes, we propose: (1) Mandatory Screening: CCHD cases should have HPLC screening for hemoglobinopathies done at diagnosis in high-prevalence areas [11,14]. (2) Timely Surgical Correction - Focus on performing early TOF repair to alleviate chronic hypoxia and polycythemia, which could lower the chances of sickling crises [7]. (3) Standardised Protocols: Develop universal guidelines for exchange transfusion and anti-seizure management in SCD-CCHD patients [6,14]. (4) Vasopressor Research: To evaluate proactive vasopressor use in TOF-SCD patients to optimise oxygenation [5]. (5) Public Health Measures: Promote awareness of SCD and genetic counselling in CCHD clinics [13,14].

Conclusion

This case demonstrates the disastrous combination of TOF and SCD, which is fuelled by thrombosis and sickling brought on by hypoxia in a 14-month-old girl. Particularly in high-prevalence areas like India, this example emphasises the vital necessity of frequent hemoglobinopathy screening in children with cyanotic congenital heart disease. Thrombotic risks could be reduced by proactive use of vasopressors, optimised exchange transfusion protocols, early TOF correction, and routine hemoglobinopathy screening. To improve results, public health activities and customised practices are needed in settings with limited resources. Consensus recommendations for this uncommon dualdisease should be established, and the efficacy of proactive vasopressor use should be validated in future research.

Author’s Disclosures

Declaration of Conflict of Interests

The authors declare that there are **no conflicts of interest** regarding the publication of this manuscript.

Ethical Approval

This study was conducted in accordance with the ethical principles for medical research involving human subjects as outlined in the **Declaration of Helsinki**. Ethical approval for the publication of this case report was obtained from the Institutional Ethics Committee. Written informed consent for publication was obtained from the patient’s parents/legal guardians.

Credit Author Statement

All authors contributed equally to the writing, development, and finalization of the case report.. All authors have read and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Funding Statement

This research received **no specific grant** from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Statement

All data supporting the findings of this study are included within the article. Additional clinical details are available from the corresponding author upon reasonable request, subject to institutional and ethical considerations.

Acknowledgements

The authors thank the patient’s family for their cooperation and consent for publication of this case report.

References

1. Moodie DS. Diagnosis and management of congenital heart disease in the adult. Cardiol Rev. 2001;9(5):276-281. Doi: 10.1097/00045415-200109000-00007. PMID: 11520451.
2. Iannucci GJ, Adisa OA, Oster ME, mcconnell M, Mahle WT. Sickle Cell Disease with Cyanotic Congenital Heart Disease: Long-Term Outcomes in 5 Children. Tex Heart Inst J. 2016;43(6):509-513. Doi: 10.14503/THIJ-15-5610. PMID: 28100970; PMCID: PMC5179156.
3. Kocis KC, Meliones JN. Cardiopulmonary interactions in children with congenital heart disease: physiology and clinical correlates. Prog Pediatr Cardiol. 2000;11(3):203-210. Doi: 10.1016/s1058-9813(00)00051-5. PMID: 10978713.
4. Shaddy RE, Penny DJ, Feltes TF, editors. Moss & Adams’ Heart Disease in Infants, Children, and Adolescents, Including the Fetus and Young Adult. 10th ed. Philadelphia: Lippincott Williams & Wilkins; 2022. Available from: https://cardiology.lwwhealthlibrary.com/book.aspx?bookid=3084§ionid=0 (accessed: 18/02/2026).
5. Kothari SS. Mechanism of cyanotic spells in tetralogy of Fallot--the missing link? Int J Cardiol. 1992;37(1):1-5. doi: 10.1016/0167-5273(92)90125-m. PMID: 1428277.
6. Adams RJ, Brambilla D; Optimizing Primary Stroke Preven-

tion in Sickle Cell Anemia (STOP 2) Trial Investigators. Discontinuing prophylactic transfusions used to prevent stroke in sickle cell disease. N Engl J Med. 2005;353(26):2769-2778. Doi: 10.1056/nejmoa050460. PMID: 16382063.

7. Fraser CD Jr, mckenzie ED, Cooley DA. Tetralogy of Fallot: surgical management individualized to the patient. Ann Thorac Surg. 2001;71(5):1556-1561. Doi: 10.1016/s0003-4975(01)02475-4. PMID: 11383800.
8. Serjeant, G.R. and Serjeant, B.E. (2001) Sickle Cell Disease. 3rd Edition, Oxford University Press New York. https://doi.org/10.1046/j.1365-2141.2001.02557.x.
9. Gladwin MT, Vichinsky E. Pulmonary complications of sickle cell disease. N Engl J Med. 2008;359(21):2254-2265. Doi: 10.1056/nejmra0804411. PMID: 19020327.
10. Ware RE, de Montalembert M, Tshilolo L, Abboud MR. Sickle cell disease: progress towards combination drug therapy. Reunion. Blood. 2017;129(5):556-566. Https://doi.org/10.1016/S0140-6736(17)30193-9
11. Kato GJ, Piel FB, Reid CD, et al. Sickle cell disease. Nat Rev Dis Primers. 2018;4:18010. Doi: 10.1038/nrdp.2018.10. PMID: 29542687.
12. Vichinsky EP, Neumayr LD, Earles AN, et al. Causes and outcomes of the acute chest syndrome in sickle cell disease. N Engl J Med. 2000;342(25):1855-1865. Doi: 10.1056/NEJM200006223422502. Erratum in: N Engl J Med 2000;343(11):824. PMID: 10861320.
13. Piel FB, Steinberg MH, Rees DC. Sickle cell disease. N Engl J Med. 2017;376(16):1561-1573. Doi: 10.1056/nejmra1510865. PMID: 28423290.
14. Yawn BP, Buchanan GR, Afenyi-Annan AN, et al. Management of sickle cell disease: summary of the 2014 evidence-based report by expert panel members. JAMA. 2014;312(10):1033-1048. Doi: 10.1001/jama.2014.10517. Erratum in: JAMA. 2014;312(18):1932. Erratum in: JAMA. 2015;313(7):729. PMID: 25203083.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Case Report

Anti-N-methyl-D-aspartate receptor autoimmune encephalitis due to ovarian teratoma in a 17-year-old young woman: case report and review of the literature

Ennio Polilli[1,*], Anna D'Andreagiovanni[2], Pierluigi Tocco[3], Maurizio Rosati[4], Ylenia Della Rocca[1,5], Giancarlo Di Iorio[1], Maria Vittoria De Angelis[3], Giustino Parruti[6], Antonella Frattari[7]

[1]*Clinical Pathology Unit, Pescara General Hospital, Pescara, Italy*
[2]*Neuro-physiopathology Unit, Pescara General Hospital, Pescara, Italy*
[3]*Stroke Unit and Emergency Neurology, Pescara General Hospital, Pescara, Italy*
[4]*Obstetrics and Gynecology Unit, Pescara General Hospital, Pescara, Italy*
[5]*Department of Life Sciences, Health and Health Professions; Link Cmpus University, Roma, Italy*
[6]*Infectious Diseases Unit, Pescara General Hospital, Pescara, Italy*
[7]*Intensive Care Unit, Pescara General Hospital, Pescara, Italy*

Article Info

[*]***Corresponding Author:***
Ennio Polilli
Clinical Pathology Unit, Pescara General Hospital, Pescara, Italy
Via Fonte Romana, 8, 65124 Pescara PE
Phone: +39 0854252598
E-mail: en.polilli@gmail.com

Keywords

Anti-NMDA receptor encephalitis, Anti-NMDAR autoimmune disease, Ovarian teratoma, Autoimmune encephalitis, Anti-NMDAR antibodies

Abstract

Many cases of encephalitis are caused by an autoimmune response, and some are associated with underlying malignancies. Here we present the case of a 17-year-old female patient who was admitted to the pediatric ward with acute gastroenteritis, seizures, and behavioral changes. A brain computed tomography (CT) scan ruled out acute infarction, haemorrhage, mass lesions, and other intracranial injuries, while the initial EEG showed no epileptic discharges. In the absence of a definitive diagnosis, midazolam was administered to treat tonic–clonic seizures, while ceftriaxone and acyclovir were started for suspected infection. However, her condition deteriorated further and she was transferred to the ICU, where she was intubated. A Cell-Based Assay (CBA) revealed the presence of anti-NMDAR antibodies and a contrast-enhanced brain MRI showed leptomeningeal enhancement without limbic involvement. She was subsequently treated with intravenous immunoglobulins and high-dose corticosteroids. A subsequent transabdominal ultrasound scan revealed an oval, fluid-filled ovarian cyst. Based on these findings, her right ovary was removed, resulting in a gradual improvement in her condition. This case highlights that paraneoplastic autoimmune encephalitis may be misclassified due to persistent psychiatric symptoms with negative EEG and CSF findings. When clinical suspicion remains high despite negative initial instrumental and laboratory findings, it is essential to repeat these investigations. In such cases, neuroantibody testing may play a crucial role in establishing the diagnosis.

Background

Encephalitis occurs in approximately 0.07–12.6 cases per 100,000 people each year, with several cases attributed to an autoimmune response [1]. The most frequently reported cases among individuals testing positive for anti-surface antibodies are anti-leucine-rich glioma-inactivated 1 (anti-LGI1) and anti-N-methyl-D-aspartate receptor (anti-NMDAR) encephalitis [1]. Anti-NMDAR autoimmune disease is caused by high levels of antibodies against the NR1 subunit of the N-methyl-D-aspartate receptor glutamate-gated ion channels and occurs more frequently in children and adult women aged 15–40 years [2].

Up to 90% of patients with anti-NMDAR encephalitis develop psychiatric or behavioral symptoms within two weeks of disease onset, leading clinicians to suspect primary psychiatric disorder [3,4,5]. These symptoms frequently co-occur with neurological signs such as impaired alertness, seizures, memory deficits, autonomic instability, speech dysfunction, decreased consciousness, and involuntary movements. Autonomic manifestations such as hypersalivation, hyperthermia, tachycardia/bradycardia, urinary incontinence, hypo/hypertension, and erectile dysfunction have also been reported [4]. Failure to screen for an underlying autoimmune neurological disease in these cases may result in misdiagnosis or delayed diagnosis [5]. Consequently, central hypoventilation and epileptic status requiring intensive care unit transfer may develop [2]. Finally, several reports have pointed out that anti-NMDAR autoimmune disease may be associated with an underlying cancer in a consistent number of cases, suggesting that it may involve paraneoplastic syndrome [6].

Herein, we describe a case of a 17-year-old woman who was admitted to our hospital with rapidly deteriorating neurological symptoms.

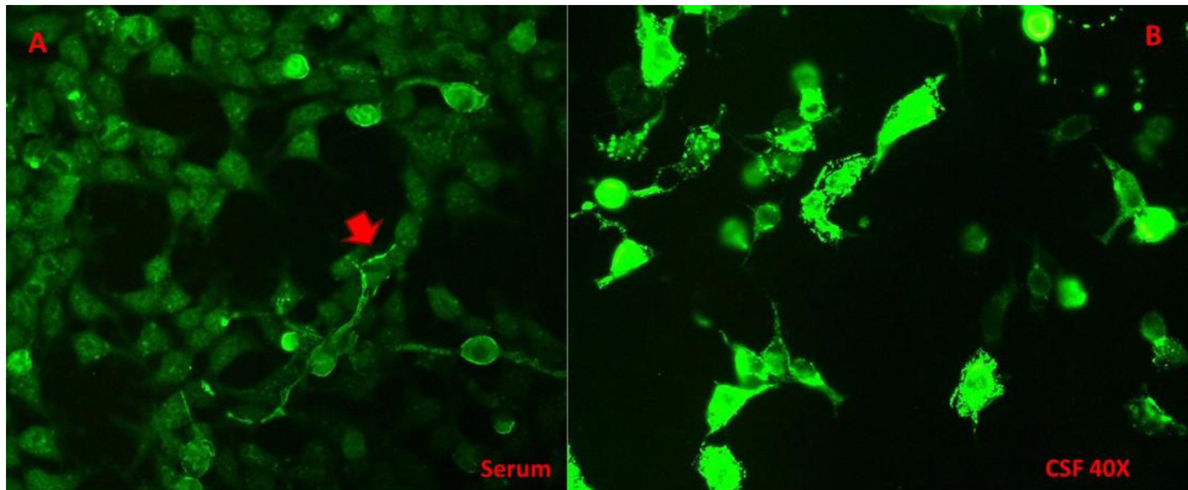
Case presentation

A 17-year-old Caucasian woman with a previous unremarkable medical history was hospitalized for acute gastroenteritis with seizures, psychosis, and personality changes in the paediatric ward of local hospital in March 2022. A brain computed tomography (CT) excluded acute brain infarction, haemorrhage, mass lesion, or other intracranial injuries. The patient showedself-limiting, stereotyped seizures associated with hyperventilation, tachycardia, and loss of consciousness. Haematochemical analyses revealed high White Blood Cell (WBC) counts (17.02 x10^3 µL), C-Reactive Protein (CRP) at 27 mg/L and unremarkable Procalcitonin (PCT) levels (<0.02 ng/mL). Cerebrospinal fluid (CSF) sampling was performed, using PCR for mycobacteria, herpes viruses and parvoviruses, yielding all negative results. CSF cultures were negative, too. In the same day, on suspicion of a conversion disorder, she was referred for consultation with a psychiatrist, who ordered an electroencephalogram (EEG) and a brain magnetic resonance imaging (MRI) scan. The EEG with intermittent light stimulation, hyperpnea, and placebo infusion documented the absence of epileptic discharges. In the absence of a clear clinical diagnosis, midazolam (0.034 mg/kg/h) was administered to treat tonic-clonic seizures, and ceftriaxone (4 g/day) and acyclovir (1,950 mg/day) for suspected infection.

Due to respiratory acidosis and an oxygen saturation (sO_2) at 82.7%, she was transferred to our ICU, where she underwent sedation and mechanical ventilation. Laboratory analyses at entry showed WBC and CRP values of 16,500 cells/µL and 68 mg/L, respectively, with PCT 0.4 ng/mL. Based on epileptogenic frontal activity during her second EEG, carbamazepine (400 mg/24 hours) was added to midazolam.

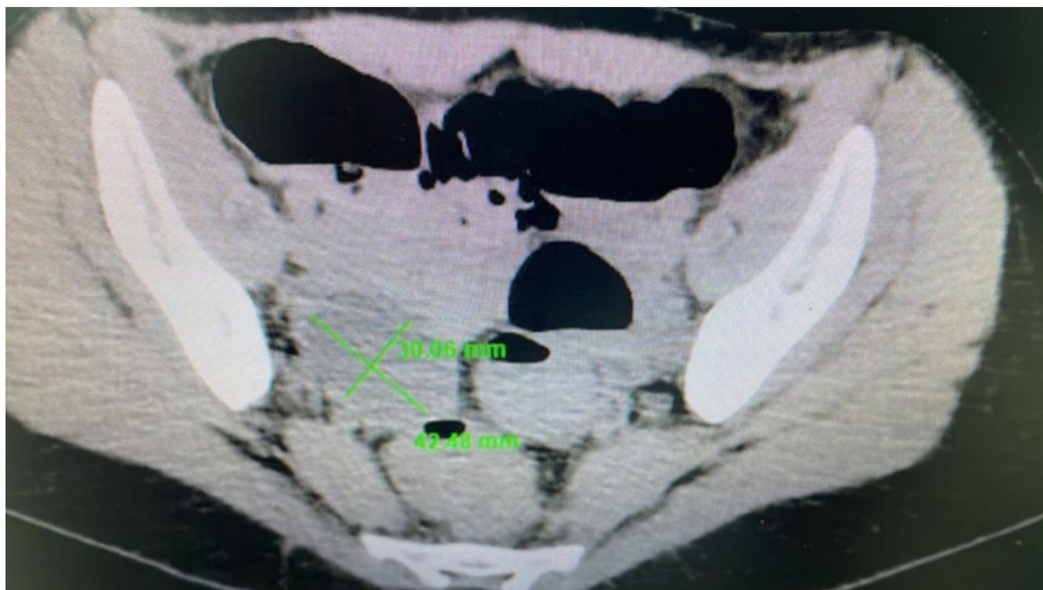
In the following day, a molecular search for herpes viruses, once more performed on peripheral blood samples, turned out negative. A second lumbar puncture, to investigate the possible presence of an autoimmune disease, revealed antibodies against neuronal surface antigens using indirect immunofluorescence assays on transfected cells (CBA, Cell-Based Assay, Euroimmun, Lubecca, Germany). Anti-NMDAR autoantibodies were analysed on the same day and were detected in both serum (Figure 1A) and CSF (Figure 1B), while anti-AMPA1/2, CASPR2, LGI1, GABArB1/2, and DPPX antibodies were all negative. Serum and CSF samples were analyzed at a 1:10 dilution and undiluted, respectively, as indicated in the manufacturer's instructions for use (IFU). The following day, the contrast-enhanced brain MRI showed no focal abnormalities on FLAIR sequences, with only leptomeningeal enhancement likely related to the previous lumbar puncture. No evidence of limbic involvement was observed.

Figure 1: Anti-NMDAR antibodies positive in serum (A) and CSF (B).



Based on these findings, the patient started high-dose methylprednisolone (1g/die) and intravenous immunoglobulin (IVIG) at 0.4 g/kg for five days, followed by IVIG alone at 0.4 g/kg daily for a further five days. On the same day, atransabdominal ultrasound showed an oval, fluid-filled cystic, 44x28 mm, right ovary mass, confirmed by a contrast abdominal/pelvic CT scan (Figure 2).

Figure 2: Fluido-cystic ovaric mass observed by abdomen/pelvic CT-scan (44x28mm).



Due to epileptic abnormalities on EEG, levetiracetam 500 mg twice daily was added. In the following days, clinical and instrumental signs of frontal seizures disappeared, sedation was tapered and the patient was extubated. On the ninth day after admission, the patient underwent laparoscopic excision of the right ovary, in parallel with biopsy of the left ovary. Right ovary histology confirmed the presence of a differentiated ovarian teratoma, with mature glial cells. Postoperative clinical conditions were stable, and the patient was discharged to the Gynaecology ward. Control NMDAR antibody assays performed on both serum and CSF after one year both showed negative results.

Discussion and conclusions

After first line investigations performed to rule out encephalitis, our patient was referred to a psychiatrist, which may have increased the risk of delayed diagnosis. There are three main reasons why our patient may have been at a greater risk for a delayed diagnosis. First, her first EEG turned out negative, in line with previous reports indicating the frequency of normal EEGs at presentation as approximately 4%. Typical EEGs show non-specific, slow, and disorganized activity, as well as electrographic seizures [7]. Second, the results of the physical examination, cell count, and chemical analysis of the first CSF sample were negative.

NMDAR disease is generally associated with inflammatory changes in the CSF, such as pleocytosis, elevated CSF protein levels, and positive OCBs (Oligoclonal bands) [8]. Third, MRI showed isolated leptomeningeal enhancement rather

thanhippocampal lesions, the most specific abnormalities associated with autoimmune encephalitis [9]. In these cases, MRI may show T2/FLAIR hyperintensities in the medial temporal lobes and occasionally in other cortical or subcortical regions. However, the scan's sensitivity and specificity are low to moderate, reliably identifying only 30–50% of patients with this condition. Many patients will have normal results, particularly in the early stages. Therefore, diagnosis should be based primarily on clinical features and confirmed by imaging and laboratory investigations of both CSF and serum [7].

Finally, in our case, a definitive diagnosis was made only after a cell-based assay, which identified the presence of anti-NMDAR antibodies in both serum and CSF.

CSF may provide more accurate information on immunopathological processes in the central nervous system and enhances the accuracy of diagnoses [10–12]. For instance, approximately 14% of patients tested positive for anti-NMDAR antibodies in CSF may have negative results in the serum [12]. Additionally, serum tests can be influenced by immune activity elsewhere in the body or by cross-reactivity, which may lead to false positives and affect treatment decisions. Furthermore, a small proportion of healthy people may test positive for NMDAR antibodies in their serum [13].

Antibody testing plays an important role in autoimmune encephalitis, as it links the clinical symptoms to their underlying cause. Detecting neuronal autoantibodies in serum and cerebrospinal fluid leads to a specific diagnosis and a prompt administration of appropriate immunotherapy. Furthermore, antibody testing may confirm the autoimmune disease etiology and support more precise classification, as specific autoantibodies are associated with distinct prognoses, therapeutic approaches, and long-term management strategies. For instance, the risk of relapse is higher in anti-LGI1 encephalitis than in anti-NMDAR encephalitis, often requiring long-term immunosuppressive therapy, whereas NMDAR encephalitis typically requires more intensive immunotherapy during the acute phase. In contrast, more intense immunotherapy is required for NMDAR encephalitis in the acute phase [14, 15, 16, 17]. Additionally, specific antibodies, such as anti-Hu or anti-Yo, may target screening for underlying tumors, aiding the identification of paraneoplastic syndromes and enabling timely cancer treatment [18]. However, a negative result does not rule out the diagnosis, and test sensitivity can vary depending on the specific antibody and type of sample. Therefore, research of antibodies should complement, rather than replace, clinical judgment, neuroimaging, electroencephalography, and cerebrospinal fluid analysis [12].

Several authors have reported that delayed initiation of immunotherapy is an important negative prognostic factor, highlighting the importance of timely and accurate diagnosis [19]. Titulaer et al. (2013) reported that early immunotherapy and avoiding ICU transfer are both independent predictors of favorable outcome [14]. Although large randomized controlled trials enrolling patients with NMDAR encephalitis are lacking, retrospective studies suggest that early diagnosis and treatment are associated with a good prognosis, fewer relapses, and milder signs and symptoms [5]. However, a consistent percentage ofpatients may require long-term care in the ICU, with increased risk of severe nosocomial pneumonia and/or other nosocomial complications [20]. Our patient was transferred to ICU due to hypoventilation, and developed a Serratia marcescens respiratory infection following intubation, promptly controlled with targeted antibiotic treatment. Thus, our case suggests that investigations for autoimmune encephalitis should be started as soon as possible, even at the Emergency Department, in patients presenting with psychiatric behaviors associated with neurological signs of encephalitis [15], as overall mortality rates for patients with NMDAR autoimmune brain disease are approximately 10%, rising to 15% for anti-NMDAR, anti-LGI1, or anti-GABABR encephalitis [16]. Patients with NMDAR encephalitis and ovarian teratomas frequently have more severe neurological signs at presentation. In such cases, early surgical treatment is mandatory for full recovery without recurrences [21]. An observational study reported that immunotherapy and tumor removal, when applicable, resulted in significant neurological improvement in 81% of patients with a median follow-up of 24 months [9]. To investigate the possible presence of an underlying teratoma, our patient underwent a CT scan, revealing an oval fluid-cystic ovarian mass in the right ovary. Timely identification of a teratoma and its subsequent resection are associated with a significantly better prognosis, including faster neurological recovery and a reduced risk of long-term complications. Most cases of mature ovarian teratoma in young patients are asymptomatic and are typically diagnosed only when abdominal pain occurs due to complications such as torsion or rupture, which may constitute surgical emergencies [22]. In some cases, the tumor is identified following a diagnosis of autoimmune encephalitis [22]. Anti-NMDAR encephalitis is the most common form of autoimmune encephalitis associated with ovarian teratomas, although this is not invariably the case [21]. The prevalence of an underlying ovarian neoplasm in these patients ranges from 26.9% to 38% [21]. Although there are currently no guidelines recommending imaging screening for teratomas in young females, imaging is essential when patients present with abdominal pain, increased abdominal girth, or neurological symptoms [22, 23]. In clinical practice, a combination of imaging techniques is typically employed to screen for an underlying tumor. Pelvic ultrasound is often the preferred option, as it is widely available, safe, and relatively inexpensive. However, it can miss smaller or more complex lesions. Therefore, in these cases, CT or MRI is the choice for a more accurate diagnosis. Additionally, if CT or MRI are also negative, but clinical suspicion remains high, PET can help identify lesions with higher metabolic activity that might otherwise be overlooked [23]. Relapse occurs in approximately 10–25% of patients with anti-NMDAR encephalitis, especially in those who do not start immunotherapy or whose underlying remains unidentified [24]. Relapses can occur months or even years after the first episode, so long-term follow-up is needed [24]. This includes regular neurological investigations, monitoring for any psychiatric or cognitive changes, and periodic tumor screening. Diagnosing relapse early allows for timely re-treatment and is linked to better outcomes [14, 24].

In line with previous reports, immunotherapy and early excision of the tumor led to rapid control of neurological signs; one month after discharge, our patient was indeed completely asymptomatic [21]. In conclusion, anti-NMDAR encephalitis is an autoimmune process that can mimic primary psychiatric illness at presentation [5]. Correct and timely identification of patients with anti-NMDAR autoimmune encephalitis remains challenging, and clinical suspicion must be put forth despite negative first-line instrumental or laboratory findings at presentation.

List of abbreviations

NMDAR: N-methyl-D-aspartate receptors
ICU: Intensive Care Unit
CSF: Cerebrospinal fluid
EEG: Electroencephalogram
MRI: Magnetic Resonance Imaging
CT: Computed Tomography
WBC: White Blood Cells
CRP: C-Reactive Protein
PCT: Procalcitonin
FLAIR: Fluid Attenuated Inversion Recovery
AMPA1/2:
Alpha-amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid
CASPR2: Contactin-associated protein-like 2
LGI1: Leucine-rich glioma inactivated 1
GABArB1/2: Subunit B1 and B2 of gamma-aminobuturic acid receptor
DPPX: Dipeptidyl-peptidase-like protein 6
ED: Emergency Department

Declarations

Ethics approval and consent to participate

All procedures performed in this study were in accordance with the ethical standards of the Health Administrative Board of the Pescara General Hospital and with the amended Helsinki Declaration. Our patient's parents provided written informed consent to use and publish anonymised clinical data for institutional research purposes upon admission to the ED.

Consent for publication

Written informed consent was obtained from patient's parents for publication of this case report and any accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Availability of data and materials

Data are available from the author upon reasonable request and with permission of Health Direction of Pescara General Hospital.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was required for this article.

Acknowledgements

The authors are indebted with all the staff of the Gynaecology and Intensive Care Units that took care of the patient during her stay in the Pescara General Hospital

References

1. Granerod J, Tam CC, Crowcroft NS, Davies NW, Borchert M, Thomas SL. Challenge of the unknown. A systematic review of acute encephalitis in non-outbreak situations. Neurology. 2010;75(10):924-932. https://doi.org/10.1212/WNL.0b013e3181f11d65.
2. Zhang Y, Liu G, Jiang M, Chen W, He Y, Su Y. Clinical Characteristics and Prognosis of Severe Anti-N-methyl-D-aspartate Receptor Encephalitis Patients. Neurocrit Care. 2018;29(2):264–272. https://doi.org/10.1007/s12028-018-0536-6
3. Kayser MS, Titulaer MJ, Gresa-Arribas N, Dalmau J. Frequency and characteristics of isolated psychiatric episodes in anti–N-methyl-d-aspartate receptor encephalitis. JAMA Neurol. 2013;70(9):1133–1139. https://doi.org/10.1001/jamaneurol.2013.3216
4. Nguyen L, Wang C. Anti-NMDA Receptor Autoimmune Encephalitis: Diagnosis and Management Strategies. Int J Gen Med. 2023;16:7–21. https://doi.org/10.2147/IJGM.S397429
5. Valente T, Gonçalves G, Duarte V, Jesus G. Anti-N-Methyl-D-Aspartate Receptor Encephalitis, A Diagnosis Not To Miss. Eur J Case Rep Intern Med. 2023;10(8):003974. https://doi.org/10.12890/2023_003974. eCollection 2023.
6. Akanuma H, Iizuka T, Abe D, et al. Paraneoplastic anti-N-methyl-D-aspartate receptor encephalitis associated with small cell lung cancer and cytotoxic T-cell-mediated pathology: Case report. Front Immunol. 2022;13:952868. https://doi.org/10.3389/fimmu.2022.952868
7. Dalmau J, Lancaster E, Martinez-Hernandez E, Rosenfeld MR, Balice-Gordon R. Clinical experience and laboratory investigations in patients with anti-NMDAR encephalitis. Lancet Neurol. 2011:63-74. https://doi.org/10.1016/S1474-4422(10)70253-2.
8. Blinder T, Lewerenz J. Cerebrospinal Fluid Findings in Patients With Autoimmune Encephalitis-A Systematic Analysis. Front Neurol. 2019;10:804. https://doi.org/10.3389/fneur.2019.00804.
9. Beatty CW, Creutzfeldt CJ, Davis AP, Hoffer Z, Khot SP. The diagnostic conundrum and treatment dilemma of a patient with a rapidly progressive encephalopathy. Neurohospitalist. 2014;4(1):34-41. https://doi.org/10.1177/1941874413496792.
10. Graus F, Titulaer MJ, Balu R, et al. A clinical approach to diagnosis of autoimmune encephalitis. Lancet Neurol. 2016;15(4):391–404.https://doi.org/10.1016/S1474-4422(15)00401-9
11. Ho ACC, Mohammad SS, Pillai SC, et al. High sensitivity and specificity in proposed clinical diagnostic criteria for anti-N-methyl-D-aspartate receptor encephalitis. Dev Med Child Neurol. 2017;59(12):1256–1260. https://doi.org/10.1111/dmcn.13579
12. Lee SK, Lee ST. The Laboratory Diagnosis of Autoimmune Encephalitis. J Epilepsy Res. 2016;6(2):45–52. https://doi.org/10.14581/jer.16010

13. Steiner J, Teegen B, Schiltz K, Bernstein HG, Stoecker W, Bogerts B. Prevalence of N-methyl-D-aspartate receptor autoantibodies in the peripheral blood: healthy control samples revisited. JAMA Psychiatry. 2014;71(7):838-839. https://doi.org/ doi:10.1001/jamapsychiatry.2014.469
14. Titulaer MJ, McCracken L, Gabilondo I, et al. Treatment and prognostic factors for long-term outcome in patients with anti-N-Methyl-D-Aspartate (NMDA) receptor encephalitis: acohort study. Lancet Neurol. 2013;12(2):157–165. https://doi.org/10.1016/S1474-4422(12)70310-1
15. Macher S, Zimprich F, De Simoni D, Höftberger R, Rommer PS. Management of Autoimmune Encephalitis: An Observational Monocentric Study of 38 Patients. Front Immunol. 2018;9:2708. https://doi.org/10.3389/fimmu.2018.02708
16. Zhong R, Chen Q, Zhang X, Zhang H, Lin W. Risk Factors for Mortality in Anti-NMDAR, Anti-LGI1, and Anti-GABABR Encephalitis. Front Immunol. 2022;13:845365. https://doi.org/10.3389/fimmu.2022.845365
17. Varley JA, Strippel C, Handel A, Irani SR. Autoimmune encephalitis: recent clinical and biological advances. J Neurol. 202;270(8):4118-4131. https://doi.org/10.1007/s00415-023-11685-3.
18. Linnoila JJ. Paraneoplastic antibodies targeting intracellular antigens. Handb Clin Neurol. 2024;200:335-346. https://doi.org/10.1016/B978-0-12-823912-4.00021-9
19. Broadley J, Seneviratne U, Beech P, et al. Prognosticating autoimmune encephalitis: A systematic review. J Autoimmun. 2019;96:24–34. https://doi.org/10.1016/j.jaut.2018.10.014
20. Niederman MS. Pneumonia: Considerations for the Critically Ill Patient. Crit Care Med. 2008;867–883. https://doi.org/10.1016/B978-032304841-5.50045-5
21. Dai Y, Zhang J, Ren H, et al. Surgical outcomes in patients with anti-N-methyl D-aspartate recep tor encephalitis with ovarian teratoma. Am J Obstet Gynecol. 2019;221(5):485.e1-485.e10. https://doi.org/10.1016/j.ajog.2019.05.026
22. O'Neill KE, Cooper AR. The approach to ovarian dermoids in adolescents and young women. J Pediatr Adolesc Gynecol. 2011;24(3):176-180. https://doi.org/10.1016/j.jpag.2010.11.006.
23. Zhang W. Ovarian teratoma in anti-NMDAR encephalitis: a life saved by imaging. Intensive Care Med. 2025. https://doi.org/10.1007/s00134-025-08180-w.
24. Ciano-Petersen NL, Villagrán-García M, Muñiz-Castrillo S, Farina A, Vogrig A, Goncalves D, Nicole F, Rogemond V, Picard G, Aubart M, Psimaras D, Oliver B, Serrano-Castro PJ, Joubert B, Honnorat J. Relapses in Anti-NMDAR Encephalitis: Clinical Characterization and Predictive Features. Neurol Neuroimmunol Neuroinflamm. 2025;12(4):e200421. https://doi.org/10.1212/NXI.0000000000200421.

Copyright© 1999–2026 International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). All rights reserved. This is a Platinum Open Access Journal distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

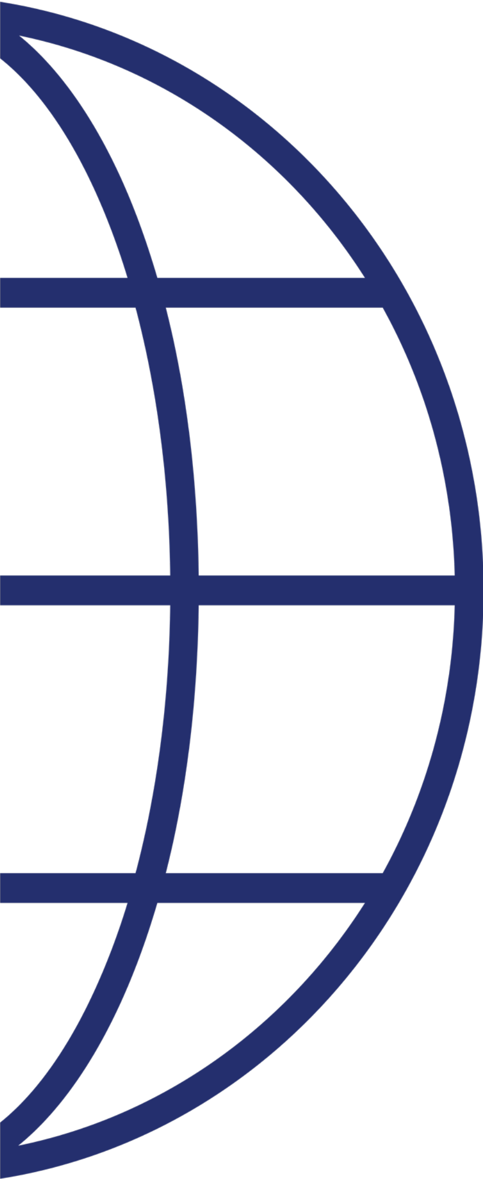
Co-Editor-in-Chief

Qing H. Meng
Department of Laboratory Medicine
The University of Texas MD Anderson
Cancer Center, Houston, USA

Kannan Vaidyanathan
Department of Biochemistry
Believers Church Medical College Hospital
Thiruvalla, Kerala, India

Editorial Board

Adil I. Khan, Temple University, Pathology and Laboratory Medicine, Philadelphia, USA
Allan S. Jaffe, Mayo Clinic, Rochester, USA
Ashish Kumar Agravatt, Biochemistry Department PDU Medical College Rajkot, Gujarat, India
Aysha Habib Khan, The Aga Khan University Hospital Main Campus Karachi: The Aga Khan University Hospital, Pathology and Laboratory Medicine, Pakistan
Béla Nagy, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Hungary
Damien Gruson, Cliniques Universitaires Saint-Luc, Laboratory Medicine, Bruxelles, Belgium
Edgard Delvin, CHU Sainte-Justine Research Center, Montréal, Québec, Canada
Ellis Jacobs, EJ Clinical Consulting, LLC, USA
Harjit Pal Bhattoa, Department of Laboratory Medicine, University of Debrecen, Hungary
Janja Marc, University of Ljubljana, Ljubljana, Slovenia
János Kappelmayer, Faculty of Medicine, University of Debrecen, Debrecen, Hungary
Joe Wiencek, Vanderbilt University School of Medicine, Pathology, Microbiology, and Immunology, USA
John Anetor, University of Ibadan, Chemical Pathology, Nigeria
Jos Wielders, Retired Head of Clinical Chemistry Lab, Netherlands
Juan Manuel Varga-Morales, Faculty of Chemical Sciences, Universidad Autónoma de San Luis Potosí, Mexico
Khosrow Adeli, The Hospital for Sick Children, University of Toronto, Canada
Lena Jafri, Aga Khan University, Pathology and Laboratory Medicine, Pakistan
Naciye Leyla Acan, Hacettepe Üniversitesi Tıp Fakültesi, Medical Biochemistry Department, Ankara, Türkiye
Nilda E. Fink, Universidad Nacional de La Plata, Argentina
Pradeep Kumar Dabla, Department of Biochemistry, G B Pant Institute of Postgraduate Medical Education & Research (GIPMER), India
Ronda Greaves, Biochemical Genetics, Victorian Clinical Genetics Services, Victoria, Australia
Sanja Stankovic, Institute of Medical Biochemistry, Clinical Center of Serbia, Belgrade, Serbia
Shanel Raghubeer, Cape Peninsula University of Technology, Biomedical Sciences, South Africa
Sibtain Ahmed, The Aga Khan University Hospital, Pathology and Laboratory Medicine, Pakistan
Stacy E. Walz, Arkansas State University, USA
Sukhes Mukherjee, Department of Biochemistry, All India Institute of Medical Sciences , India
Swarup A. V. Shah, Laboratory Medicine, PD Hinduja National Hospital and Medical Research Centre, India
Tomris Ozben, Akdeniz University, Antalya, Türkiye
Udara Dilrukshi Senarathne, University of Sri Jayewardenepura Faculty of Medical Sciences, Biochemistry, Nugegoda, Sri Lanka



Publisher: IFCC Communications and Publications Division (IFCC-CPD)

Copyright © 2026 IFCC. All rights reserved.

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.

This is a Platinum Open Access Journal distributed under the terms of the *Creative Commons Attribution Non-Commercial License* which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Published by:

www.ifcc.org