International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

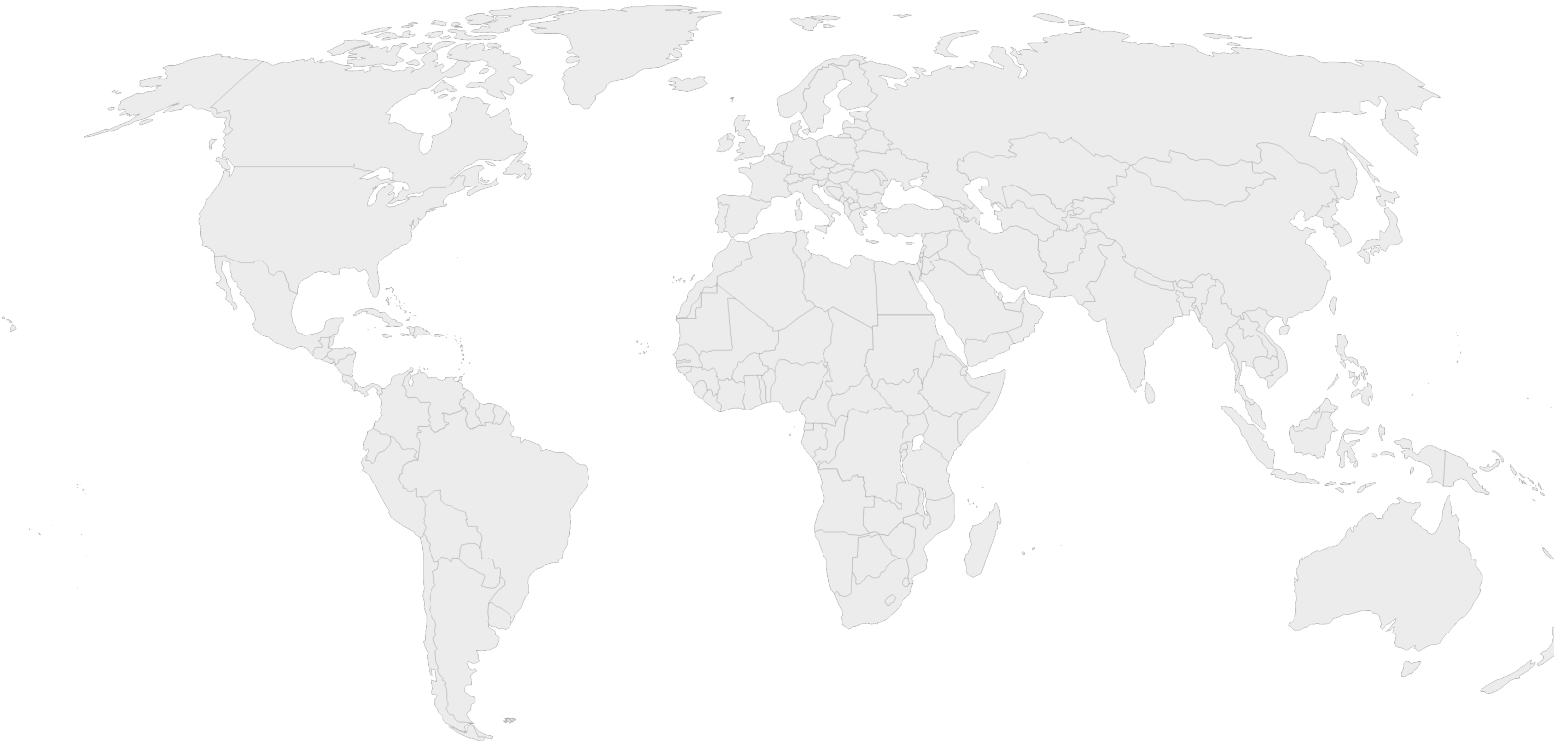
Communications and Publications Division (CPD) of the IFCC

October 2026
Volume 37
Number 5
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

Population-Specific Testosterone Reference Intervals in Pakistani Men: Application of the RefineR Algorithm

Sibtain Ahmed, Aiza Zeeshan, Imran Siddiqui 783

Ethical Use of Social Media by Lab Professionals: A Self-Assessment SAFE-POST Checklist from the IFCC Task Force on Ethics

Sibtain Ahmed, Melissa M. Budelier, Sudip Kumar Datta, Martina Zaninotto, Angela Amayo, Julian Verona, Joesph Wiencek 787

Performance comparison of two computer-aided sperm analysis systems

Efrem Bonelli, Roberta Cont, Patrizia Bartolotta, Luca Pio Dicembre, Diego Franciotta, Adriano Anesi 792

Role of Klotho Deficiency in Mediating Off-Target Effects of Elevated FGF23 in Chronic Kidney Disease: A Clinical and Molecular Correlation Study

Subha Sree S, Swathy Moorthy 798

Longitudinal Laboratory and Diagnostic Profile of Patients with Prostate Cancer and PCA3 Expression: A Case Series

Divye Ahlawat, Aditya Khatri, Amanpreet Singh Gulian, Debprosit Karmakar, Sakshi Rai, Shubhangi Mukherjee, Suman Kumar Ray, Ketan Mehra, Sukhes Mukherjee 808

Association of CTLA-4 Gene Polymorphisms (CT60 and +49 A/G) with Hypothyroidism: A Case-Control Study in a Tertiary Care Hospital

KR Gopala Krishnan, Hariom Sharma, Piyush Tailor, Shailesh Patel, D Shreshtta, Hina Baraiya 819

Evaluation of Infectious Disease Immunoassays for HIV 1/2 Ab-Ag Combo and Syphilis Ab on Alinity i, Maglumi X3 and CL-1200i CLIA Systems

Diana G. Coste, Stefan Dascalu, Anthony J. Nagle 828

An Exploratory Study to assess the Feasibility and Acceptability of Clinical Infection Score Estimation in Patients with Suspected Infection at a tertiary care hospital

Prashanth Mada, Radhika Soanker, Sakthivadivel Varatharajan, Shyamala Ravikoti, Shantveer G Uppin, Tara Roshni Paul, Gaddala Mary Sowjanya 845

Evaluation of Pediatric Pleural Effusion: Insights from a Tertiary Care Center in North India

Ashna Kumar, Sugandhi Malgotra, Keshav Mehra, Preeti Sharma, Mohd. Razaq, Nikhil Rajvanshi 853

In this Issue

Evaluation of Anti-HCV, HBsAg and HAV IgM on Three Different Chemiluminescence Immunoassay Systems for Hepatitis Screening

Diana G. Coste, Stefan Dascalu, Anthony J. Nagle 862

Comparison of Cardiovascular Risk Classification by Established Risk Scores Across Glycemic Status in an Indian Population

Lathangi G, Pooja R, Nisarga T, Manish Kumar Pandey, Jadeppa G, Swetha N. Kempegowda 880

The application of indirect methods to establish and verify Argentinian paediatric thyroid function test reference intervals

Veronica E. Zaidman, Melina B. Pelanda, Noelia Dujovne, Natalia Gazek, Claudia Insua, Rodolfo E. Recalde, Florencia Zabaleta, Eduardo A. Chaler, Juan M. Lazzati 892

Multicomponent CAPA Programme Incorporating Cloud-Based Laboratory Information System Migration for Post-Analytical Error Reduction in an Accredited Stand-Alone Diagnostic Laboratory: A 13-Month Before-and-After Study

Jagdish Chandarana, Kinjal Butani 901

Circulating HMGB1 as a Potential Biomarker of Inflammation in Type 2 Diabetes Mellitus and Its Complications: A Systematic Review and Meta-Analysis

Usha Kiran Pothu, Roshan Kumar Mahat 915

When Xanthomas Tell the Story: A Child with De Novo Familial Hypercholesterolemia

Mohnish Darshan, Ajay Kumar, Pooja Soni, Bhavna Dhingra 940

Limitations of Transferrin- and UIBC-Based Estimation of Transferrin Saturation in Decompensated Chronic Liver Disease: A Case Series

Jagadish Ramasamy 945

Hyperferritinemia and Beyond: Biochemical Insights from a Case Series of Secondary Hemophagocytic lymphohistiocytosis (HLH)

Lekha Priyadharshini Kamarajan, Sushil Kumar, Ravikirti, Janmeshwar Prasad, Mala Mahto 951

Medical Laboratory Accreditation Programs: A Comparative Review of International and National Frameworks

Abdelaziz Sanad Abdelaziz, Samir Y. Marzouk, Mai S. Mabrouk 960

In this Issue

The Expanding Role of the Clinical Laboratory in Patient-Centered Care : Lessons from Sweat Testing
Rihab Makhlouf, Khouloud Mzid, Ameni Feki, Fatma Makni Ayadi 972

An International Federation of Clinical Chemistry and Laboratory Medicine Policy on Managing Conflict of Interest in Laboratory Medicine Practice Guidance Documents
Tze Ping Loh, Julian Verona, Melissa M. Budelier, Joesph R. Wiencek 978

The Human-AI Hybrid: A New Partnership for the Future of Empirical Science
Eleftherios P. Diamandis 981

Real world implementation of whole blood hemolysis indices on potassium results at the Point of Care
Nazmin Bithi, Earnest JP Daniel, Michael D. Dowlin, Brandy Recio, Sridevi Devaraj 985

Letter to the editor

Population-Specific Testosterone Reference Intervals in Pakistani Men: Application of the RefineR Algorithm

Sibtain Ahmed[1], Aiza Zeeshan[2], Imran Siddiqui[1*]

[1]*Section of Clinical Chemistry, Department of Pathology & Laboratory Medicine, the Aga Khan University Karachi, Pakistan*
[2]*Medical College, the Aga Khan University Karachi, Pakistan*

Article Info

[*]***Corresponding Author:***
Imran Siddiqui
Section of Clinical Chemistry
Department of Pathology and Laboratory Medicine
The Aga Khan University Hospital, Pakistan
Phone: + 92 21-34861927
E-mail: imran.siddiqui@aku.edu

Keywords

Testosterone, men, reference interval, indirect, Pakistan

Dear Editor,

Serum testosterone levels are essential for assessing normal physiological processes in males as well as identifying pathological conditions such as Hypogonadism [1]. Interpretation of these values relies heavily on reference intervals established for specific populations and age groups, taking cognisance of age, race, nutritional status among others due to physiological variations [2,3]. However, in Pakistan, laboratories often adopt reference intervals derived from international populations, which may not accurately represent local ethnic and physiological differences. Although reference intervals can also be established using direct methods involving healthy participants, limited adherence to guidelines from the Clinical and Laboratory Standards Institute may affect their reliability [4]. Recently, indirect statistical approaches have emerged and are supported by the International Federation of Clinical Chemistry and Laboratory Medicine as efficient methods for estimating reference intervals using existing laboratory data [5,6]. Therefore, this study utilizes patient data to derive reference intervals for serum testosterone and compares them with existing Pakistani and international values.

This study was conducted at the Section of Clinical Chemistry, Department of Pathology and Laboratory Medicine, Aga Khan University (AKU). Data mining of serum testosterone results from 2008-2025 was done on a target population ranging from 19-39 (Mean: 29.82 ± 5.32) years of age after approval from institutional ethical review committee. This age group was selected because testosterone levels typically peak in early adulthood (around 20–30 years) and begin to gradually decline from approximately 30–40 years onward, reflecting the transition from physiological maturity to age-related hormonal decline. [7,8]. Serum total testosterone concentrations were measured using electrochemiluminescence immunoassay (ECLIA) on Roche platforms throughout the study period. Testing was initially performed on the Roche Elecsys e170 analyzer, followed by the Roche Elecsys e411, and currently on the Roche Cobas e801 analyzer. All analyses were conducted according to the manufacturer's protocols in an accredited laboratory with routine internal quality control and participation in external quality assessment program to ensure analytical reliability.

The corresponding clarification has been added to the Methods sectionThe program used is the RefineR package (version 1.0.0) due to its low mean error in separating non pathological values from pathological ones and further using them to derive reference intervals [9]. The AKU clinical lab handles sample from all over the country through its network on more than 300 phlebotomy centers spread across all of Pakistan, hence our sample was representative of the geographical distribution and all ethnicities of the country. Due to the retrospective nature of the dataset, pre-analytical variables such as timing of sample collection (e.g., morning vs non-morning samples) and fasting status could not be standardized. In cases where multiple test request forms existed for the same individual, only the results from the initial sample were considered for the final analysis. 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 was obtained from the institutional review committee of Aga Khan University (ERC#2025-11907-36009), and all data were fully anonymized prior to analysis.

Given the right-skewed distribution of testosterone values, a Box–Cox transformation was applied to approximate a Gaussian distribution as required by the RefineR algorithm. The algorithm iteratively models the central distribution corresponding to the non-pathological population. Using the Box-Cox transformation method, the 2.5th percentile and 97.5th percentile were used as the lower and upper reference limit with a confidence interval of 95% as shown in Figure 1 was estimated for 75,547 values. No explicit exclusion of pathological values was performed; instead, the RefineR algorithm statistically models the central non-pathological distribution within the dataset, thereby minimizing the influence of outliers and pathological results. From this dataset, the estimated reference interval for serum testosterone in Pakistani males aged 19–39 years was 166–942 ng/dL (5.76–32.69 nmol/L). The distribution of values demonstrated a right-skewed pattern with a clear central peak representing the presumed healthy population, from which the algorithm derived the reference limits. The calculated reference interval was subsequently compared with previously reported international reference ranges in Table 1.

Figure 1: Estimated Reference Intervals (from Patients' Samples).

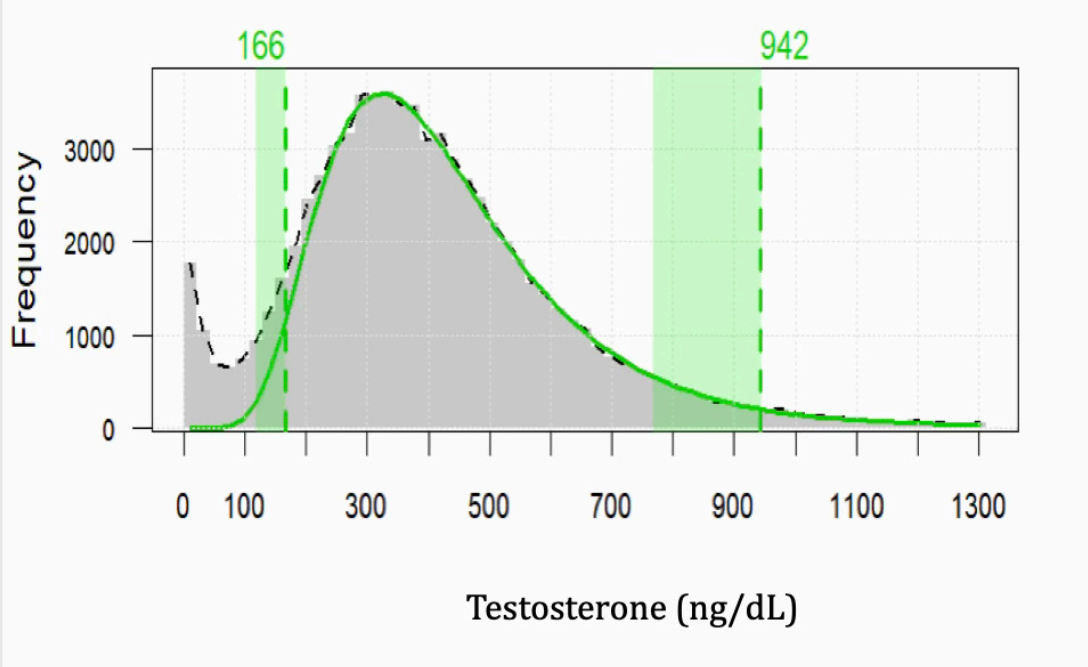


Table 1: Testosterone reference intervals reported in published studies.

Author	Study Type	Population	Age (years)	Calculation method	Testosterone reference Intervals (ng/dL)
Joeri Walravens et al, 2025[10]	Cohort	Belgium	18–39	Direct	184–749
Xue Qin et al, 2012[11]	Cohort	China	20–69	Direct	363–1066
Myrna BuenaluzSedurante et al, 2020[12]	Cohort	Philippines	21–40	Direct	211.4–1528.2
Ravi Jasuja et al, 2022[13]	Cohort	US	19–39	Direct	152.5–925
Thomas G Travison et al, 2017[14]	Cohort	US and Europe	19–39	Direct	264–916
Pruktanakul T et al, 2021[15]	Cohort	Thailand	18–65	Direct	211–970 ng/dL

In conclusion, this study establishes testosterone RIs for Pakistani males using an indirect statistical approach. The estimated reference interval in our study shows both overlap and notable differences when compared with previously published international ranges. For example, studies from European and US cohorts report lower limits ranging from approximately 152 to 264 ng/dL and upper limits between 749 and 925 ng/dL, while data from Asian populations demonstrate wider variability. In comparison, our findings demonstrate a relatively broader upper limit and comparable lower limit, which may reflect population-specific biological variation, differences in assay methods, and the use of an indirect statistical approach.

These observations highlight that while international RIs provide a useful framework, locally derived, data-driven estimates may better reflect the underlying population distribution. However, given the indirect nature of our methodology, these intervals should be interpreted cautiously and require validation using well-characterized healthy cohorts before routine clinical implementation.

Declaration of Conflict of interest
None.

Ethical Approval
Ethical approval was obtained from the institutional review committee, and all data were fully anonymized prior to analysis.

Author's Disclosures
None.

Author Statements
Sibtain Ahmed contributed to the conceptualization of the study, manuscript drafting, and overall supervision. Aiza Zeeshan contributed to data analysis and manuscript writing. Imran Siddiqui contributed to the study concept, critical revision of the manuscript, and intellectual input. All authors reviewed and approved the final version of the manuscript.

Funding
No funding received.

Data Availability Statements
Available on reasonable request from corresponding author.

References

1. Mooradian AD, Morley JE, Korenman SG. Biological actions of androgens. Endocr Rev. 1987;8(1):1–28. doi:10.1210/edrv-8-1-1.
2. Sikaris KA. Physiology and its importance for reference intervals. Clin Biochem Rev. 2014;35(1):3–14.
3. Heald AH, Ivison F, Anderson SG, Cruickshank K, Laing I, Gibson JM. Significant ethnic variation in total and free testosterone concentration. Clin Endocrinol (Oxf). 2003;58(3):262–266. doi:10.1046/j.1365-2265.2003.01653.x.
4. Fleming JK, Katayev A, Moorer CM, Ward-Jeffries DA, Terrell CL. Development of nation-wide reference intervals using an indirect method and harmonized assays. Clin Biochem. 2022;99:20–59. doi:10.1016/j.clinbiochem.2021.10.001.
5. Doyle K, Bunch DR. Reference intervals: past, present, and future. Critical Reviews in Clinical Laboratory Sciences. 2023;60(6):466-482. Doi: https://doi.org/10.1080/10408363.2023.2196746
6. Jones GRD, Haeckel R, Loh TP, Sikaris K, Streichert T, Katayev A, et al. Indirect methods for reference interval determination – review and recommendations. Clin Chem Lab Med. 2019;57(1):20–29. doi:10.1515/cclm-2018-0073.

7. Banica T, Verroken C, Reyns T, Mahmoud A, T'Sjoen G, Fiers T, et al. Early decline of androgen levels in healthy adult men: an effect of aging per se? A prospective cohort study. J Clin Endocrinol Metab. 2021;106(4):e1074–e1083. doi:10.1210/clinem/dgaa915.
8. Narinx N, Marriott RJ, Murray K, Adams RJ, Ballantyne CM, Bauer DC, et al. Sociodemographic, lifestyle, and medical factors associated with calculated free testosterone concentrations in men: individual participant data meta-analyses. Eur J Endocrinol. 2024;191(5):523–534. doi:10.1093/ejendo/lvae133.
9. Ammer T, Schützenmeister A, Prokosch HU, Rauh M, Rank CM, Zierk J. refineR: a novel algorithm for reference interval estimation from real-world data. Sci Rep. 2021;11:16023. doi:10.1038/s41598-021-95301-2.
10. Walravens J, Snaterse G, Narinx N, Reyns T, Van Uytfanghe K, Vanderschueren D, et al. Age-stratified reference ranges for directly measured serum free testosterone in community-dwelling and healthy men. J Clin Endocrinol Metab. 2026;111(3):e787–e793. doi:10.1210/clinem/dgaf507.
11. Qin X, Lv H, Mo Z, Chen Z, Lin L, Peng T, et al. Reference intervals for serum sex hormones in Han Chinese adult men from the Fangchenggang Area Male Health and Examination Survey. Clin Lab. 2012;58(3–4):281–290.
12. Buenaluz-Sedurante M, Co MIK, Dagang DJT, Bruno RG, Sarmiento AJN, Tee ML. Reference intervals of total testosterone in adult Filipino men. Int J Endocrinol. 2020;2020:8877261. doi:10.1155/2020/8877261.
13. Jasuja R, Pencina KM, Spencer DJ, Peng L, Privat F, Dhillo W, et al. Reference intervals for free testosterone in adult men measured using a standardized equilibrium dialysis procedure. Andrology. 2023;11(1):125–133. doi:10.1111/andr.13310.
14. Travison TG, Vesper HW, Orwoll E, Wu F, Kaufman JM, Wang Y, 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(4):1161–1173. doi:10.1210/jc.2016-2935.
15. Pruktanakul T, Taemkaew K, Liabsuetrakul T, Leelawattana R, Kietsiriroje N. Normal Reference Ranges of Serum Testosterone and Gonadotropins in Thai Fertile Men: A Cross-sectional Study in a Single Tertiary Center. Journal of Health Science and Medical Research. 2021;39(2):123-133. Doi: http://dx.doi.org/10.31584/jhsmr.2020766

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.

Letter to the editor

Ethical Use of Social Media by Lab Professionals: A Self-Assessment SAFE-POST Checklist from the IFCC Task Force on Ethics

Sibtain Ahmed[1], Melissa M. Budelier[2, 3], Sudip Kumar Datta[4], Martina Zaninotto[5], Angela Amayo[6], Julian Verona[7], Joesph Wiencek[8*]

[1]*Department of Pathology and Laboratory Medicine, Aga Khan University, Karachi, Pakistan*
[2]*TriCore Reference Laboratories, Albuquerque, NM, USA*
[3]*Department of Pathology, University of New Mexico School of Medicine, Albuquerque, NM, USA*
[4]*Department of Laboratory Medicine, All India Institute of Medical Sciences, New Delhi, India*
[5]*QI.LAB.MED, spin-off of the University, Padova, Italy*
[6]*Department of Human Pathology, Faculty of Health Sciences, University of Nairobi*
[7]*Servicio de Laboratorio, Hospital de Balcarce "Dr. Felipe A. Fossati", Buenos Aires, Argentina*
[8]*Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA*

Article Info

[*]***Corresponding Author:***
Joesph R. Wiencek
VUMC Department of Pathology, Microbiology, and Immunology
1301 Medical Center Drive, 4605 TVC, Nashville, TN 37232-5310
Telephone: 615-322-3532
Fax: 615-343-9563
E-mail: joe.wiencek@vumc.org

Keywords

Social Media, Post, Ethics

To the Editor,

Social media has been a game changer for the mass dissemination of scientific knowledge in the rapidly advancing field of laboratory medicine. Platforms like LinkedIn, Facebook, X (previously Twitter), WhatsApp, Telegram, YouTube, and professional online forums have evolved into useful avenues for case discussions, professional networking, continuing education, and scientific exchange in recent years. Pathologists and clinical laboratory professionals, who have conventionally been in less visible roles, are now increasingly projecting via social media to advocate for laboratory medicine, expand outreach, and foster interdisciplinary collaborations [1]. This evolution became more evident during the COVID-19 pandemic era, where laboratory medicine's online presence via social media platforms was widely portrayed not only for public health awareness, but also for the dissemination of scientific information [2].

Social media landscape extensively covers the attributes of big data i.e., volume, velocity, variety, veracity, and value. Every day, vast amounts of content are uploaded across diverse platforms and reach broad audiences. One on hand, it offers important windows for communication, education, and knowledge exchange. Conversely, it also poses challenges regarding the accuracy, interpretation, and enduring impact of information that can spread rapidly and remain publicly accessible long after it is shared. [3]. It has the potential to contribute to misinformation, particularly when complex laboratory concepts are generalized, preliminary findings are discussed without adequate context, or emerging evidence is circulated before sufficient validation [4]. For instance, posts describing novel laboratory tests or preliminary research findings prior to sufficient clinical validation may inadvertently overestimate their utility. Similarly, presenting rare case experiences without mentioning limitations may lead to misconceptions. For laboratory professionals, "out of context" may include isolated screenshots of case discussions, excerpts from scientific debates, or educational posts viewed without accompanying explanations, disclaimers, or supporting evidence.

Many educational posts that include lab images, diagnostic reports, screenshots, case discussions, or rare case studies have the potential of inadvertently compromising patient confidentiality, even when identifiers have been removed [5]. In some cases, even minor details such as metadata, internal institutional information, partially hidden accession numbers, very rare laboratory findings, or supplementing contextual clinical information may still allow indirect patient identification. This in turn poses valid reputational risks for individuals, laboratories, and institutions if it portrays, inappropriate communication, or confidentiality breach which has the potential to lead to weakened public trust [6]. Likewise, respectful communication is equally important in online spaces, where scientific discussions may naturally involve differing opinions. In situations such as responding to an opposing viewpoint, correcting misinformation, laboratory professionals should strive to maintain a respectful, constructive, and professional tone. Even when disagreements arise, comments should remain focused on the scientific issue rather than becoming dismissive. In various instances, professionalism extends beyond clinical or patient based discussions. Lab professionals should keep in mind that posts made in a professional capacity or publicly associated with their identity may also affect public confidence in their affiliated institution and the profession.

Data privacy and consent are also vital concerns in the use of mass messaging tools. Platforms such as WhatsApp are increasingly used for rapid professional communication, educational case discussions, troubleshooting, and peer consultation. While these applications offer considerable convenience, they also raise important ethical challenges [7]. Professional WhatsApp groups deserve careful consideration where informal discussions, and screenshots can be re-shared without permission, and even supposedly de-identified laboratory data could unintentionally allow incidental patient identification. From beneficence perspective, these tools should complement rather than replace standard institutional communication systems and secure clinical pathways [8]. Furthermore, as artificial intelligence (AI) tools become increasingly utilized for content creation, there is growing emphasis to remain vigilant and strict verification of potential AI generated inaccuracies, fabricated references, and illustrative images that could unintentionally misrepresent laboratory findings.

Keeping in view the core ethical principles (Beneficence, Maleficence, Autonomy and Justice) and scientific integrity, laboratory professionals have a responsibility to ensure that posting via their social media platforms is accurate, balanced, and presented within its proper scientific context [9, 10]. Caution is needed when discussing emerging evidence, such as pre-validation interpretation or overstating the strength of findings has the potential to lead to more harm than the intended benefit. With this objective in sight, the concept of an "Ethics Checklist for Social Media" was first proposed at the IFCC Task Force on Ethics (TF-E) meeting in Milan in 2026. The preliminary draft was prepared by a full member of the IFCC TF-E who had previous experience in social media coordination with the IFCC Committee on Internet and Digital Communications (C-CIDC). The checklist was built upon fundamental ethical principles that guide professional practice in laboratory medicine. Its domains were focused on encouraging beneficial and responsible communication while emphasizing respect for individuals, protection of confidentiality, scientific integrity, transparency, accountability, and the maintenance of public trust.

The draft of the checklist was subsequently presented to the IFCC Task Force on Ethics for discussion and review. Seven members provided feedback on the wording, clarity, ethical relevance, applicability across different laboratory medicine scenarios, and feasibility of use as depicted in Figure 1.

The overall idea was balancing comprehensiveness with practicality, ensuring that the checklist addressed emerging ethical challenges while remaining relatively simple and user friendly. Based on the feedback received via an iterative expert consensus process, the checklist was revised. The revised checklist was circulated to all Task Force members, and the final version was approved by consensus before submission. (Table 1), comprising experienced laboratory professionals from diverse countries, and practice settings. Although formal external validation with additional stakeholder groups was beyond the scope of this initial work, future studies may evaluate its usability.

The checklist has the potential for the development across diverse laboratory settings and can be relevant for laboratory organizations, scientific societies, and individual professionals seeking to practice ethical social media engagement. This check

list is not meant to be used as a formal clearance process, but rather a useful self-assessment tool to promote introspection prior to posting. When any item on the checklist gets a negative response, users should edit or postpone it and, if necessary (for example, if there are issues with confidentiality, institutional policy, or legal inferences), ask for feedback from an impartial and trusted peer before publishing. Institutions may also integrate the checklist into social media policies, orientation programs or professionalism training. The TF-E acknowledge that the checklist relies on professional judgment and should complement, rather than replace, institutional governance frameworks, in accordance with relevant local legislation, regulatory requirements, and institutional policies.

Figure 1: Flow chart of checklist development process.

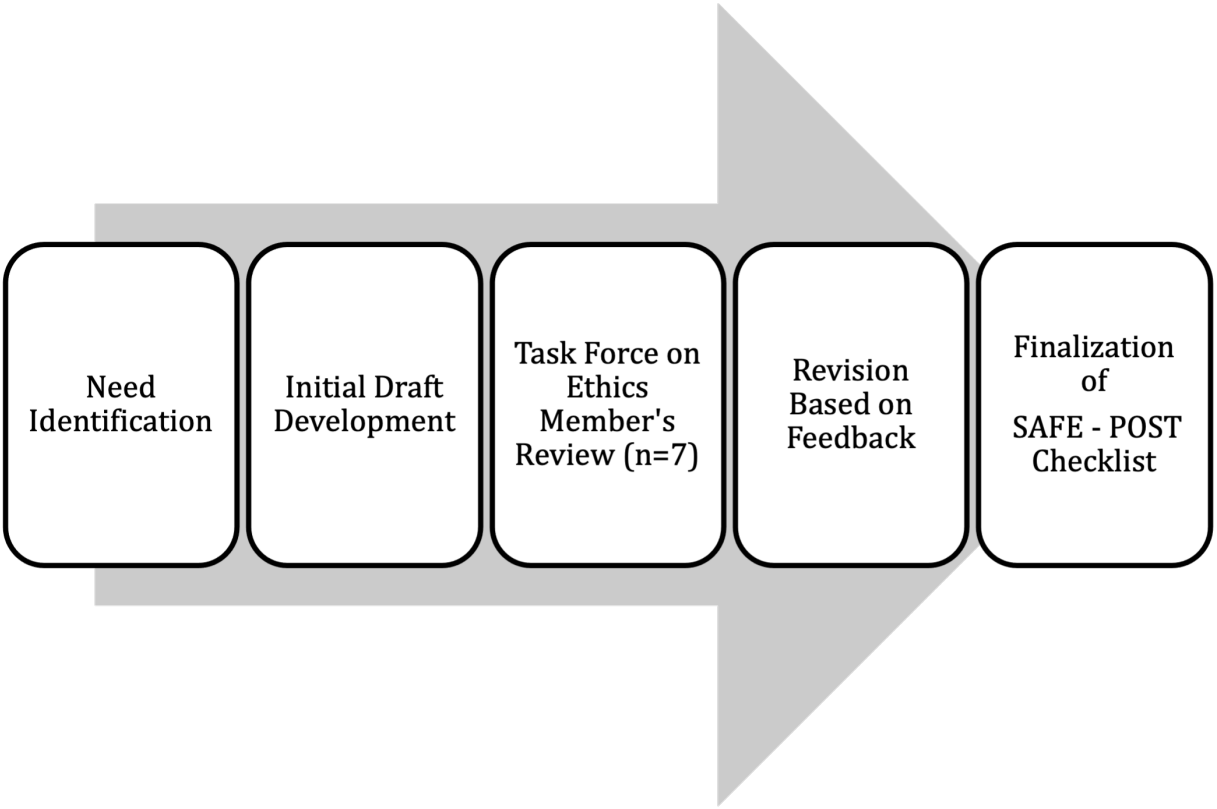


Table 1: SAFE-POST Ethical Checklist.

Domain	SubDomain	SelfAssessment	Response Check
Autonomy	Transparency	Have conflicts of interest, affiliations, and sponsorships been appropriately disclosed?	□ Yes □ No If No: modify before posting
Autonomy	Professionalism	Would this post reflect positively on the profession and the institution if reviewed publicly?	□ Yes □ No If No: modify before posting
Autonomy	Respect for Persons	Does the post maintain respectful language toward peers, patients, and differing opinions?	□ Yes □ No If No: modify before posting
Beneficence	Educational Benefit	Does this post provide meaningful educational, scientific, or public health value?	□ Yes □ No If No: modify before posting
Beneficence	Public Trust	Does this post strengthen trust in laboratory medicine and scientific communication?	□ Yes □ No If No: modify before posting

Domain	SubDomain	SelfAssessment	Response Check
Beneficence	Scientific Integrity	Is the content evidence- based, appropriately contextualized, and free from misleading claims or exaggeration? If AI assisted content generation was used, has the information been reviewed for scientific accuracy?	□ Yes □ No If No: modify before posting
Nonmaleficence	Harm Prevention	Could this post contribute to misinformation, unsafe clinical behavior, panic, or misinterpretation?	□ Yes □ No If Yes: modify before posting
Nonmaleficence	Confidentiality & Privacy	Have patient identifiers, metadata, images, and contextual clues been adequately removed or consent obtained?	□ Yes □ No If No: modify before posting
Nonmaleficence	Accountability	Would this post remain acceptable if permanently archived or publicly scrutinized or reshared out of context?	□ Yes □ No If No: modify before posting
Justice	Professional Competence & Equity	Is this post within area/scope of professional competence, acknowledging limitations and promoting unbiased scientific engagement?	□ Yes □ No If No: modify before posting
Justice	Acknowledgment & Attribution	If the post is reshared/forwarded/reposted has the original sender been acknowledged?	□ Yes □ No If No: modify before posting

Disclosures of Interest

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

Ethical Approval

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

Funding

None.

Data Availability

Not applicable.

Author Statement

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

References

1. Casey P., Schukow Adam L., Booth Kamran M. Mirza, Ryan Philip Jajosky; #PathTwitter: A Positive Platform Where Medical Students Can Engage the Pathology Community. Arch Pathol Lab Med 1 February 2023; 147(2):135–136. doi: 10.5858/arpa.2022-0282-ED. PMID: 36453869
2. Zehra Mz M, Un Nisa Zn Z, Ahmed Sa S. From Bench to Bytes: Utility of Social Media Platforms by Pathology Residents. EJIFCC. 2023;34(4):317-324. PMID: 38303748; PMCID: PMC10828540.
3. Denniss E, Lindberg R. Social media and the spread of misinformation: infectious and a threat to public health. Health Promot Int. 2025;40(2):daaf023. doi: 10.1093/heapro/daaf023. PMID: 40159949; PMCID: PMC11955583.
4. Feng X, Luo J, Yang Y, El Baz D, Shi L. Health Misinformation Detection: Approaches, Challenges and Opportunities. Inquiry. 2025;62:469580251384784. doi: 10.1177/00469580251384784. Epub 2025 Nov 4. PMID: 41189452; PMCID: PMC12589804
5. Gardner JM, Allen TC. Keep Calm and Tweet On: Legal and Ethical Considerations for Pathologists Using Social Media. Arch Pathol Lab Med. 2019;143(1):75-80. doi: 10.5858/arpa.2018-0313-SA. Epub 2018 Aug 22. PMID: 30132683.
6. Garmon EH, Morris KC, McAllister RK. Preserving medical professionalism in the age of social media. JCA Advances. 2024;1(3-4):100055. Doi: https://doi.org/10.1016/j.jcadva.2024.100055
7. Mars M, Scott RE. WhatsApp in clinical practice: a literature review. The Promise of New Technologies in

an Age of New Health Challenges. 2016:82-90. PMID: 27782019

8. Johnston MJ, King D, Arora S, Behar N, Athanasiou T, Sevdalis N, Darzi A. Smartphones let surgeons know WhatsApp: an analysis of communication in emergency surgical teams. The American Journal of Surgery. 2015;209(1):45-51. DOI: 10.1016/j.amjsurg.2014.08.030. PMID: 25454952
9. Gardner J.M., McKee P.H. Social media use for pathologists of all ages. Arch Pathol Lab Med. 2019;143(3):282–286. doi: 10.5858/arpa.2018-0431-ED. PMID: 30816833
10. Flippo A, Dixit B, Schukow CP, Hacking SM, Song L, Fiock K, Golab K, Sonawane S, Alter DN, Rohde RE, Baskota SU. Social media in pathology and laboratory medicine: a systematic review. Academic Pathology. 2024;11(4):100151. PMID: 39559455. DOI: 10.1016/j.acpath.2024.100151

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

Performance comparison of two computer-aided sperm analysis systems

Efrem Bonelli[1], Roberta Cont[1], Patrizia Bartolotta[2], Luca Pio Dicembre[1], Diego Franciotta[1*], Adriano Anesi[1]

[1]*UOM, Laboratory of Clinical Pathology, Department of Laboratories, Asuit, Santa Chiara Hospital, Largo Medaglie d'Oro 9, Trento, Italy*
[2]*Department of Clinical and Evaluative Epidemiology, Trento Health Service (Asuit), Viale Verona, Trento, Italy*

Article Info

[*]***Corresponding Author:***
Diego Franciotta
Laboratory of Clinical Pathology, Department of Laboratories, APSS of Trento, Largo Medaglie d'Oro 9, 38123 Trento, Italy
E-mail: diego.franciotta@apss.tn.it

Keywords

Computer assisted semen analyzer, Male infertility, Semen analysis

Abstract

Background: Automated systems are promising tools for the standardization of laboratory semen analysis, but the assessments on these systems are still insufficient to support their use in daily practice. We aimed to evaluate the analytic performance of two computer-aided sperm analysis (CASA) systems, having manual evaluation as a reference.

Methods: Fifty-two seminal fluid samples were analyzed in accordance with the WHO 6th Edition guidelines at the Clinical Pathology Laboratory of Trento. A LensHooke X1 PRO instrument and a SCA SCOPE system were used for the automated analysis of the following parameters: sperm concentration count, total sperm motility count, spermatozoa with progressive motility count, spermatozoa with rapid progressive motility count, and spermatozoa morphology.

Results: Correlation and consistency between manual method and both the two CASA systems were good-to-moderate for all the parameters, with the exception of morphology (no correlation). However, the agreement was widely variable, with slight to moderate, or even considerable differences, and proportional biases towards overestimations.

Conclusions: Our data suggest that the two here studied CASA systems cannot be an alternative to expert manual semen analysis at present but might assist and facilitate the analysis.

Introduction

Human semen examination is the most widely used test for the diagnosis of male infertility, recommended as first laboratory analysis in couple infertility work-up, and necessary before storing, or donating semen samples [1]. Misdiagnosis, and delay in infertility treatments can follow inaccurate laboratory reports, mainly due to analytical issues [1].

Seminal fluid analysis usually relies on experienced operators' subjectivity for appearance, viscosity, and fluidization, as well as, on microscope, sperm concentration, total number, viability, and morphology [2, 3]. The lack of between-laboratory standardization, assay heterogeneity, and difficulty in identifying pathologic cut-offs still affect expert manual analysis [4].

Therefore, computer-aided sperm analysis (CASA) systems have been developed, igniting debates on their current performance and potentialities [3, 4]. Main suitable tasks include basic sperm function and motility assessment. Notwithstanding promising results [5-8], these systems are still insufficiently performing to enter routine laboratory practice [4]. Their main weakness are the lack of standardized algorithms and of extensive quality control procedures, and small volumes required for testing, which are scarcely representative of the whole content sample [4].

Having microscope analysis as the referral technique, we aimed to evaluate the analytic performance of two CASA systems in assessing semen count, mobility, and morphology.

Materials and methods

Fifty-two seminal fluid samples were from persons older than 18 years, who underwent sperm analysis at the APSS of Trento, between July and August 2024, for the following purposes: infertility issues in couples, medically assisted procreation procedures, post-vasectomy control, or sperm cryopreservation. Samples with sperm concentration counts lower than 2 million/mL were excluded from the study. The samples were collected and analyzed in accordance with the WHO Laboratory Manual for the Examination and Treatment of Seminal Fluid criteria [3]. Compliance with the pre-analytical phase rules of the WHO Manual was verified, particularly those regarding ejaculatory abstinence for 2-7 days before seminal fluid collection. The study complied with the ethical principles outlined in the 2008 revised Helsinki Declaration.

The following four key parameters were considered: a) sperm concentration count (SCC), b) total sperm motility count (TSMC), c) spermatozoa with progressive motility count (SPMC), d) spermatozoa with rapid progressive motility count (SRPMC). In addition, spermatozoa morphology was evaluated, subdivided into the following categories: normal, head defects, intermediate trait defects, tail defects.

Seminal fluids were analyzed with: a) light microscopy (manual), b) LensHooke X1 PRO instrument (Bonraybio, Taichung City, Taiwan), c) SCA SCOPE system, version 6.6 (Microptic, Barcelona, Spain).

LensHooke X1 PRO technology exploits integrated high-resolution, autofocusing optical lens in combination with an automatic computational system. Seminal fluid samples are applied into two special containment wells and then analyzed. SCA SCOPE is equipped with an optical microscope with high-resolution lenses, and with an integrated high-resolution digital camera with a Complementary Metal Oxide Semiconductor sensor. Four slides are set up for each sample for microscopy evaluation. Images, obtained from different areas of the slide, are displayed on a monitor along with the test results.

Automated semen analysis was performed in accordance with the manufacturers' protocols. Seminal fluid samples were simultaneously analyzed with both the systems, and by experienced operators (E.B. and R.C.) under a microscope.

Statistical analysis

Bland-Altman analysis was used to assess the agreement between manual sperm analysis and the two automated systems. Mean difference (bias) and 95% limits of agreement (LoA) were calculated to determine the range within which most differences fell. Using a two-way random-effect model for absolute agreement, we assessed the between-measurement degree of agreement through intraclass correlation coefficients (ICC; 95% confidence interval), which were interpreted as follows: <0.50, poor; 0.50-0.74, moderate; 0.75-0.90, good; >0.90, excellent. Statistical analysis was performed with "Software R" (4.3.3 version; R foundation, Vienna, Austria) [9].

Results

Table 1 shows the data on seminal fluid analysis obtained with the two CASA technologies and with microscope evaluation.

Table 1: Data of the seminal fluid analysis obtained with two CASA technologies and with microscope evaluation.

	Microscopy evaluation	LensHooke	SCA SCOPE
SCC, 10^6/mL, median (IQR)	19 (4, 44)	23 (6, 65)	25 (11, 51)
TSMC, %, median (IQR)	51 (37, 64)	55 (25, 68)	37 (20, 54)
SPMC, %, median (IQR)	36 (21, 50)	38 (19, 58)	25 (13, 36)
SRPMC, %, median (IQR)	21 (14, 36)	16 (5, 22)	10 (5, 21)
Morphology			
Normal, %, median (IQR)	4.0 (2.0, 6.0)	4.0 (3.0, 5.0)	1.5 (0.9, 3.0)
Head defects, %, median (IQR)	85 (80, 91)	80 (76, 85)	93 (83, 97)
Intermediate trait defects, %, median (IQR)	31 (30, 40)	50 (44, 57)	85 (78, 90)
Tail defects, %, median (IQR)	17 (13, 25)	25 (23, 26)	0 (0, 0)

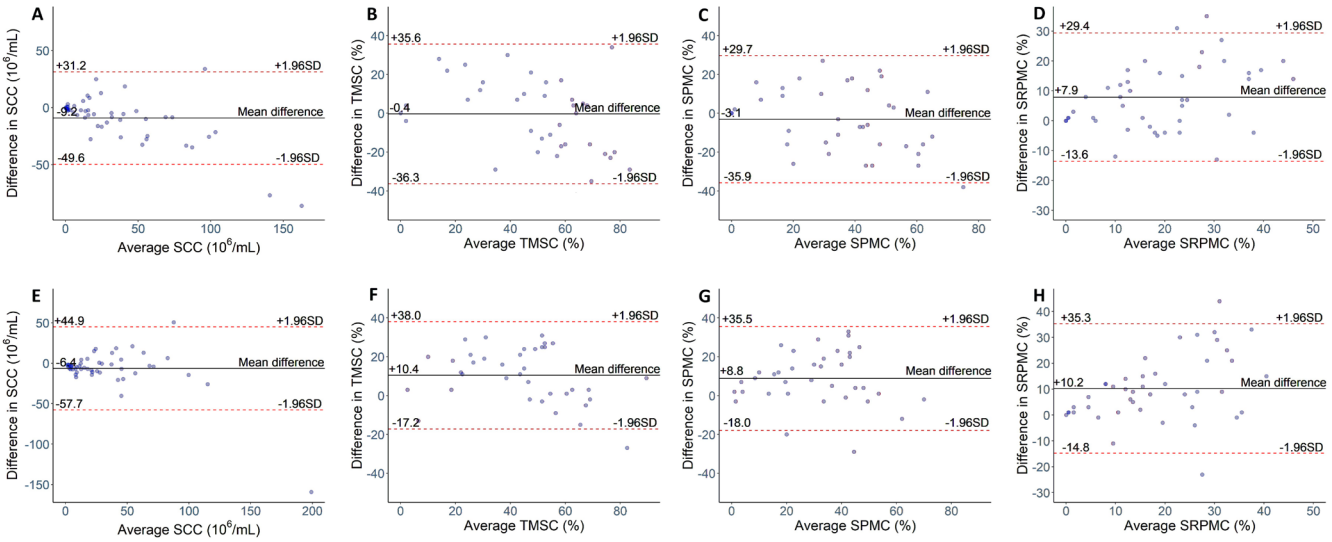
CASA, computer-aided sperm analysis; SCC, sperm concentration count; TSMC, total sperm motility count; SPMC, spermatozoa with progressive motility count; SRPMC, spermatozoa with rapid progressive motility count; IQR, interquartile ranges

Correlations between microscope measurements vs CASA results were classifiable as: a) "good" for SSC [ICC, 0.842 (IQR, 0.701-0.914; p<0.001) vs LensHooke, and 0.767 (0.621-0.862; p<0.001) vs SCA SCOPE]; b) "moderate" for TSMC [0.721 (0.520-0.846; p<0.001) vs LensHooke, and 0.712 (0.311-0.870; p<0.001) vs SCA SCOPE]; c) "moderate" for SPMC [0.694 (0.483-0.829; p<0.001) vs LensHooke, and 0.66 (0.311-0.832; p<0.001) vs SCA SCOPE]; d) "moderate" for SRPMC [0.58 (0.189-0.783; p=0.003] vs LensHooke, and "poor" for SRPMC [ICC, 0.406 (0.025-0.662; p=0.018) vs SCA SCOPE]. CASA technologies and microscope evaluation did not correlate for spermatozoa morphology.

LensHooke X1 PRO vs microscope evaluation

For SCC, between-method mean difference was -9.2 × 10^6/mL [95% LoA range, -49.6×10^6/mL to +31.2×10^6/mL] (Figure 1 panel A), indicating LensHooke X1 PRO underestimations. Most data points cluster around the mean difference value, and dispersion increases at the highest SCC, with few outliers. This suggests that between-method discrepancies became more pronounced as SCC increased, potentially indicating a proportional bias. For TMSC, the overall agreement was good, with minimal systematic bias (-0.4%). However, the wide variability (95% LoA range, -36.3% to +35.6%) suggests substantial differences in outliers (Figure 1 panel B). A slight proportional bias could be present, as larger deviations tended to associate with higher TMSC values. Agreement for SPMC was strong, with a small systematic bias (-3.1%). However, variability was considerable, with 95% LoA ranging from -35.9% to +29.7% (Figure, 1 panel C). For SRPMC, LensHooke X1 PRO overestimated the counts (Figure 1 panel D). Moderate variability was observed (95% LoA: -13.6% to +29.4%), with a possible proportional bias, since the discrepancies increased proportionally to higher SRPMC values.

Figure 1: Difference against average data between manual vs LensHooke X1 PRO (A-D), and manual vs SCA SCOPE (E-H) for the following parameters: sperm concentration count (SCC), total sperm motility count (TSMC), spermatozoa with progressive motility count (SPMC), spermatozoa with rapid progressive motility (SRPMC); 95% limits of agreement (broken lines).



SCA SCOPE vs microscope evaluation

A mean difference of -6.4×10⁶/mL was found for SCC (Figure, panel E), indicating that SCA SCOPE slightly underestimated sperm concentration. LoA extended from +44.9×10⁶/mL to -57.7×10⁶/mL, highlighting considerable variability. For TMSC, mean difference was +10.4% (95% LoA range, -17.2% to +38.0%; Figure, panel F), suggesting a general overestimation and moderate variability of the parameter. A systematic bias of +8.8% affected SPMC, along with wide variability (95% LoA range, -18.0% to +35.5%; Figure, panel G). For SRPMC, SCA SCOPE overestimated the values by 10.2% (Figure, panel H), with moderate variability (95% LoA range, -14.6% to +35.3%). The data distribution suggested a possible bias proportional to higher SRPMC values.

SCA SCOPE vs LensHooke X1 PRO evaluation

Overall, the performance of the two CASA systems did not differ significantly. Correlations between SCA SCOPE vs LensHooke results were classifiable as: a) "good" for SSC [ICC, 0.895 (IQR, 0.822-0.939; p<0.001)]; b) "moderate" for TSMC [0.509 (0.270-0.690; p<0.001); c) "moderate" for SPMC [0.528 (0.247-0.717; p<0.001); d) "moderate" for SRPMC [0.712 (0.543-0.827; p<0.001].

Discussion

According to the WHO, seminal fluid analysis is the reference test for defining the role of male infertility in couple infertility. Instruments that can ensure greater standardization with good performance in this setting could improve laboratory practice, especially in non-specialist laboratories. CASA systems are expected to perform in the basic sperm analysis [4], which is why we selected sperm concentration, motility, and morphology for this study.

For sperm concentration, we found good correlation and consistency between manual method and the two CASA systems. However, it is well-known that at low concentrations only the human eye can discriminate all the components that make spermatozoa identifiable as such [3]. This is a difficult to overcome technological limit at present. However, both CASA systems extend the 'validated' concentration ranges below the threshold of 2 million spermatozoa/ml, which we, on the contrary, have identified as the lower limit of accuracy. Correlation and consistency were only moderate when the motility parameters were compared between microscopy visualization and the CASA systems.

Our findings on agreement between manual method and the CASA systems showed wide variability, ranging from slight to, more frequently, moderate-to-considerable differences for all the considered parameters. The biases were often towards overestimations, and affected both the systems, with slight differences.

Efficient use of CASA technologies in seminal fluid analysis is still a matter of debate [2, 5-8, 10-12]. Agarwal and colleagues, who developed the LensHooke X1 Pro-based system, found good correlation data on sperm concentrations and motility in comparison with manual microscopy [2]. Analogously, these parameters resulted well correlated in a comparison study between an earlier version of the SCA SCOPE and manual evaluation [5]. Both these studies, however, did not take into account rapid and slow progressive motility, namely, the parameters reintroduced in the sixth edition of the WHO3 manual [3].

In a more recent study, Jani and colleagues compared microscopy visualization and LensHooke X1 Pro system, showing good correlation for sperm concentration and for total motility [8]. As for the agreement, they judged it acceptable

and concluded that LensHooke X1 Pro system could help standardize reports in non-specialist laboratories [8].
Overall, both our and literature data suggest that evidence is still insufficient to make CASA completely substitute semen manual analysis in routine laboratory practice [4]. In particular, our results cannot confirm the high degree of agreement with the manual method regarding sperm concentration and motility, which the SCA SCOPE and LensHooke X1 Pro systems are supposed to guarantee, in accordance with the manufacturers' claims. These technologies could help assess sperm motility, especially fast and slow progressive motility, since the human eye does not have standard parameters to define these features [2, 5, 11-12]. However, they are at risk of mistaking debris for sperm at low concentrations, thus falsely lowering the motility values, and of failing in recognizing spermatozoa at high concentrations [4]. Practically, the use of microscopy should be mandatory to assign motility around and below the decision-making limits [3].
At present, CASA technologies could thus be useful in the initial evaluation of seminal fluids, necessitating expert human eye confirmation, especially when sperm concentration and motility are low. As for morphology, artificial intelligence is expected to improve CASA performance.
In conclusion, the analytic performance of LensHooke x1 Pro and SCA SCOPE was substantially comparable. LensHooke X1 Pro instrument has the advantage of 'all-in-one' analysis, which takes few seconds and less staff time. SCA SCOPE is more complex, since slide preparation needs time and dexterity, and morphological analysis time is longer. These features make SCA SCOPE more suitable for research and second-level examination, whereas LensHooke x1 Pro could better fit for laboratories dealing with large number of samples to be processed as soon as possible.

Author contributions
Conceptualization: Efrem Bonelli, Adriano Anesi; Material preparation, data collection: Efrem Bonelli, Roberta Cont; Statistical analysis: Patrizia Bartolotta; Original draft preparation: Efrem Bonelli, Diego Franciotta; Review, editing, and final approval: Efrem Bonelli, Roberta Cont, Patrizia Bartolotta, Luca Pio Dicembre, Diego Franciotta, Adriano Anesi

Data availability
The data supporting the findings of this study are available on request from the corresponding author.

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

Ethical Approval
The local Institutional Review Board deemed the study exempt from review.

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

References
1. Vander Borght M, Wyns C. Fertility and infertility: Definition and epidemiology. Clin Biochem 2018;62:2-10. doi:10.1016/j.clinbiochem.2018.03.012
2. Agarwal A, Henkel R, Huang CC, Lee MS. Automation of human semen analysis using a novel artificial intelligence optical microscopic technology. Andrologia 2019;51:e13440. doi:10.1111/and.13440.
3. WHO. WHO laboratory manual for the examination and processing of human semen. 6th edition, 2021. Geneva, Switzerland; World Health Organization 2021.
4. Baldi E, Gallagher MT, Krasnyak S, Kirkman-Brown J, Editorial Board Members of the WHO Laboratory Manual for the Examination and Processing of Human Semen. Extended semen examinations in the sixth edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen: contributing to the understanding of the function of the male reproductive system. Fertil Steril 2022;117:252-257. doi:10.1016/j.fertnstert.2021.11.034
5. Lammers J, Splingart C, Barriere P, Jean M, Freour T. Double-blind prospective study comparing two automated sperm analyzers versus manual semen assessment. J Assist Reprod Genet 2014;31:35-43. doi:10.1007/s10815-013-0139-2
6. Schubert B, Badiou M, Force A. Computer-aided sperm analysis, the new key player in routine sperm assessment. Andrologia 2019;51:e13417. doi:10.1111/and.13417
7. Lammers J, Chtourou S, Reignier A, Loubersac S, Barriere P, Freour T. Comparison of two automated sperm analyzers using 2 different detection methods versus manual semen assessment. J Gynecol Obstet Hum Reprod 2021;50:102084. doi:10.1016/j.jogoh.2021.102084
8. Jani E, Bozzola M, Zagler EM, Roccaforte V, Daves M. Comparison between a new device for the semen quality analysis and the manual microscopic evaluation in a not specialistic clinical laboratory. Adv Lab Med 2024;5:402-406. doi:10.1515/almed-2024-0089
9. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2024. Available from: https://www.R-project.org/. Accessed, 10 October 2025.
10. Finelli R, Leisegang K, Tumallapalli S, Henkel R, Agarwal A. The validity and reliability of computer-aided semen analyzers in performing semen analysis: a systematic review. Transl Androl Urol 2021;10:3069-3079. doi:10.21037/tau-21-276
11. Talarczyk-Desole J, Berger A, Taszarek-Hauke G, Hauke J, Pawelczyk L, Jedrzejczak P. Manual vs. computer-assisted sperm analysis: can CASA replace manual assessment

of human semen in clinical practice? Ginekol Pol 2017;88:56-60. doi:10.5603/GP.a2017.0012
12. Gallagher MT, Smith DJ, Kirkman-Brown JC. CASA: tracking the past and plotting the future. Reprod Fert Develop 2018;30:867-874. doi:10.1071/RD17420

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

Role of Klotho Deficiency in Mediating Off-Target Effects of Elevated FGF23 in Chronic Kidney Disease: A Clinical and Molecular Correlation Study

Subha Sree S[1], Swathy Moorthy[1*]

[1]*Department of General Medicine, SRIHER, Porur, Chennai, Tamilnadu, India*

Article Info

Corresponding Author:
Swathy Moorthy, Professor
Department of General Medicine, SRIHER, Porur, Chennai, Tamilnadu, India
E-mail: drswathymoorthy@sriramachandra.edu.in
Phone: 91-9444016401

Keywords

FGF23, Klotho, off-target signalling, FGFR4, carotid intima-media thickness, cardiovascular remodelling, FGF23: Klotho ratio, chronic kidney disease, vascular calcification

Abstract

Background: The increase of plasma fibroblast growth factor 23 (FGF23) gives rise to pathological cardiovascular effects in chronic kidney disease (CKD) via Klotho-independent off-target receptor activation. In chronic kidney disease (CKD), soluble alpha-Klotho, an essential co-receptor for directing FGF23 signalling to kidneys declines, resulting in its stronger effect at off-target sites. We hypothesized that Klotho deficiency alters the intensity of cardiovascular remodelling induced by FGF23 in a manner that is measurable in a clinical setting.

Methods: This was a cross-sectional study carried out in 160 hospitalised patients with CKD from stages G2–G5 (40 per stage). Serum FGF23 as well as soluble α-Klotho were measured using ELISA. The cardiac restructuring was measured by means of carotid intima-media thickness (CIMT), ankle-brachial pressure index (ABPI), left ventricular hypertrophy (LVH), and aortic valve (AV) sclerosis. Patients were stratified by Klotho tertiles (<37.6, 37.6–69.4, >69.4 pg/mL) and FK (FGF23: Klotho) ratio tertiles. We computed Spearman correlations between log-FGF23 and cardiovascular outcomes in each stratum of Klotho. A comparison of cardiovascular outcomes was made from a quadrant interaction analysis across four axis states, using low/high FGF23 along with low/high Klotho.

Results: Klotho levels were not significantly different between CKD stages (median 48.1-58.2 pg/mL, Kruskal-Wallis p=0.907) and 22.5-32.5% of patients in each stage were severely deficient (<33 p g/mL). Similarly, the FGF23: Klotho ratio increased from G2 (median 1.75) to G5 (4.47; p<0.001), indicating progressive axis uncoupling. The relationship between log-FGF23 and CIMT was significantly influenced by Klotho stratum: r=0.642 (p<0.001) in the low-Klotho tertile; r=0.317 (p=0.021) in the middle tertile; and r=0.047 (p=0.739, not significant) in the high-Klotho tertile. The patients with high FK ratio had

greater CIMT than those with a low FK ratio. The results of the study show that the CIMT significantly increased across FK ratio tertiles. The group of seven patients with high FGF23 and low Klotho (mean FGF23=758 pg/mL, Klotho=23 pg/mL) had a mean CIMT of 1.364±0.069 mm, significantly greater than that of the reference group with low FGF23 and high Klotho (1.065±0.181 mm; p=0.0003). When analyzed independently, there was no significant difference in LVH and AV sclerosis across Klotho strata.

Conclusions: Klotho deficiency is an important permissive factor that enhances the carotid atherosclerotic impact of high FGF23 during CKD. This uncoupling is represented by the FGF23: Klotho ratio, which better correlates to vascular remodelling than either biomarker alone. These findings provide clinical and molecular evidence that Klotho availability halts the magnitude of off-target FGF23 effects and that measures promoting Klotho availability is preferred over FGF23 suppression.

Introduction

The key determinant of renal phosphate handling and 1,25-dihydroxyvitamin D synthesis is the Fibroblast growth factor 23 (FGF23)–α-Klotho axis. Under physiological conditions, the circulating osteocyte-derived hormone FGF23 binds to and activates a ternary complex consisting of fibroblast growth factor receptor 1 (FGFR1) and its obligate co-receptor α-Klotho, which is expressed almost exclusively in renal distal tubule [1, 2]. The Klotho-gated signalling mechanism gives specifies the actions of FGF23, ensuring that the phosphaturic and vitamin D-suppressing actions are limited to the tubular epithelium, preventing off-target actions.

The elegant architecture of the glomerulus is progressively dismantled in chronic kidney disease (CKD). At the earliest stages of renal impairment, even before the emergence of hyperphosphatemia and secondary hyperparathyroidism, levels of FGF23 start to rise exponentially. In end-stage kidney disease (ESKD) for example, they may reach over 1000-fold the normal value [3]. Meanwhile, expression of α-Klotho in the renal tubule is downregulated by uraemic toxins, inflammatory cytokines and via epigenetic changes [4, 5]. Owing to this dual dysregulation, FGF23 progressively uncouples from its renal gatekeeper, clinically quantified by the FGF23: Klotho ratio. As Klotho declines, FGF23 in concentrations that ultimately overwhelm the physiological levels commences activation of alternative, lower-affinity receptors in tissues that do not constitutively express α-Klotho. The fibroblast growth factor receptor 4 (FGFR4) on cardiomyocytes is the most important of these. FGF23, in a strictly Klotho-independent manner, activates the calcineurin–nuclear factor of activated T-cells (NFAT) hypertrophic signalling cascade in cardiomyocytes through this receptor promoting pathological left ventricular hypertrophy (LVH) [6, 7]. In the vasculature, depletion of soluble Klotho, the ectodomain shed from tubular cells into the circulation, removes its inhibition of Na^+/Pi co-transporters in vascular smooth muscle cells (VSMCs). This allows phosphate uptake, osteogenic trans differentiation, and medial calcification [5, 8]. Klotho also normally promotes eNOS via the PI3K/Akt pathway; its loss limits NO bioavailability, leading to endothelial dysfunction and accelerates subclinical atherosclerosis [9]. An important yet overlooked question is whether Klotho deficiency acts simply as a bystander, alongside FGF23 elevation, independently resulting in cardiovascular injury, or as a modifier whose absence exacerbates the cardiovascular toxicity of a raised plasma FGF23 concentration. If Klotho acts as a gate, limiting access of FGF23 to off-target receptors, then lower Klotho levels should correlate with stronger associations between FGF23 concentration and cardiac remodelling outcomes.

The current study aims at addressing this question directly among CKD patients. It therefore assessed a cross-sectional cohort of 160 CKD patients across stages G2–G5, with concurrent measurements of serum FGF23, soluble Klotho, and four parameters of cardiovascular remodelling. Through stratifying FGF23-cardiovascular correlations by Klotho status and assessing the FGF23: Klotho ratio as an index of axis uncoupling, we aimed to provide clinical and molecular evidence for Klotho gating of FGF23 mediated off-target cardiovascular effects.

Methods

Study design and participants

A cross-sectional observational study was conducted at a tertiary care facility. Adult patients who were consecutively admitted in the hospital with certified CKD as per the KDIGO 2012 criteria, having stable renal function and without any intercurrent acute kidney injury were enrolled in the study. A total of 160 participants were included, stratified a priori into four groups of 40 by CKD stage G2 (eGFR 60–89 mL/min/1.73m²), G3 (30–59), G4 (15–29), and G5 (<15 mL/min/1.73m², pre-dialysis). Those patients who underwent renal replacement therapy were excluded. The additional exclusion criteria included active malignancy, acute inflammatory disease, severe hepatic disease and pregnancy. The study employed the CKD-EPI 2021 creatinine equation [10] to estimate GFR. The study was conducted in accordance with the Declaration of Helsinki and Institutional ethics committee approval and written informed consent were obtained from all participants.

Biochemical measurements

All laboratory parameters were measured on fasting conditions: serum creatinine, blood urea nitrogen (BUN), electrolytes, sodium, potassium, chloride, and bicarbonate, haemoglobin, albumin and fasting diabetic profile. The concentration of serum FGF23 was assessed using a C-terminal FGF23 ELISA. As 1000 pg/mL was the upper limit of the assay, any readings above this were recorded as >1000 and for analyses requiring

continuous values were treated as 1000 pg/mL. The quantity of soluble alpha-Klotho was assessed by using the sandwich ELISA assay. Before the correlation analyses, the values for FGF23 were log transformed. We calculated the FGF23: Klotho ratio (FK ratio) by dividing FGF23 (pg/mL) with Klotho (pg/mL) to serve as a continuous index of axis uncoupling. For a comparison of sub-groups, FGF23 and Klotho were divided into tertiles, and the boundaries for FK ratio tertiles were T1/T2: 1.25 and T2/T3: 4.17.

Cardiovascular remodelling assessments

Four cardiovascular remodelling markers were measured. (1) CIMT- Carotid Intima Media Thickness, measured by B-mode ultrasonography, was performed on the far wall of right and left common carotid arteries (1 cm proximal to the carotid bifurcation) to obtain mean of three measurements in each side in mm. A CIMT of over 1.0 mm was regarded as subclinical atherosclerosis and over 1.2 mm as significant atherosclerosis. (2) The ankle-brachial pressure index (ABPI) was measured with a hand-held Doppler probe; ABPI <0.90 was taken as indicative of peripheral arterial disease (PAD), whilst values >1.40 were flagged as potentially non-compressible. (3) Left ventricular hypertrophy (LVH) was defined as echocardiographically left ventricular mass index >115 g/m² in men and >95 g/m² in women based on ASE criteria recorded as binary (yes/no). (4) As per the echocardiography findings, aortic valve sclerosis was identified as valve leaflet thickening or calcification without a haemodynamically significant obstruction. Thus, aortic valve sclerosis was recorded as binary.

Statistical analyses

Continuous variables were examined for distributional characteristics and reported either as mean ± standard deviation when normally distributed or as median with interquartile range when skewed. Categorical data were summarised as counts and percentages. Group comparisons across CKD stages G2–G5 were conducted using the Kruskal–Wallis test for continuous measures and the chi-square test for categorical measures, with Dunn's procedure applied for post-hoc contrasts. Because of non-normality, serum FGF23 concentrations were log-transformed prior to correlation analyses. For continuous analyses, values above the assay detection ceiling of 1000 pg/mL were truncated at that limit. All statistical procedures were implemented in Python version 3.11 using the pandas, scipy, scikit-learn, and statsmodels libraries.

Associations between components of the FGF23–Klotho axis and cardiovascular remodelling indices (carotid intima-media thickness [CIMT], ankle-brachial pressure index [ABPI], left ventricular hypertrophy [LVH], and aortic valve sclerosis) were assessed using Spearman rank correlation. Patients were further stratified by tertiles of circulating Klotho and by tertiles of the FGF23: Klotho ratio to explore the modifying influence of Klotho deficiency on FGF23-related cardiovascular changes. A composite vascular remodelling score was derived by standardising CIMT, inverting ABPI, and incorporating LVH and aortic valve sclerosis status. Interaction analyses were performed to evaluate the combined effect of high/low FGF23 and high/low Klotho status on vascular remodelling. Statistical significance was defined as $p < 0.05$.

Results

Participant characteristics and CKD staging

A total of 160 patients (mean age 59.8±12.7 years; 33.8% female) were included. 66.9% had diabetes mellitus and 65.0% had hypertension. Results stratified by CKD stage are shown in Table 1. eGFR was significantly different between groups (G2: 69.7±12.8; G3: 42.9±9.9; G4: 22.6±6.4; G5: 8.1±3.5 mL/min/1.73m²; H=142.95, p<0.001), confirming appropriate allocation. The G5 patients were significantly younger (51.7±13.1 years) than G2-G4 patients which perhaps indicates an earlier disease onset in this South Asian population or a referral bias.

Table 1: Baseline clinical, biochemical, and cardiovascular characteristics stratified by CKD stage.

Variable	CKD G2 (n=40)	CKD G3 (n=40)	CKD G4 (n=40)	CKD G5 (n=40)	p value
Age (years), mean±SD	63.5±13.0	60.6±11.1	63.5±9.8	51.7±13.1	<0.001
Female sex, n (%)	11 (28)	15 (38)	14 (35)	14 (35)	0.729
Diabetes mellitus, n (%)	25 (63)	27 (68)	33 (83)	22 (55)	0.062
Hypertension, n (%)	23 (58)	21 (53)	29 (73)	31 (78)	0.022
eGFR (mL/min/1.73m²), mean±SD	69.7±12.8	42.9±9.9	22.6±6.4	8.1±3.5	<0.001
Creatinine (mg/dL), mean±SD	1.14±0.24	1.72±0.45	2.93±0.68	10.54±19.4	<0.001
BUN (mg/dL), mean±SD	15.6±7.4	21.5±9.3	38.4±16.1	48.3±21.9	<0.001
Bicarbonate (mmol/L), mean±SD	22.4±2.7	21.3±4.6	18.5±3.1	17.9±4.2	<0.001
Haemoglobin (g/dL), mean±SD	11.7±2.4	10.2±2.0	9.4±1.7	8.6±1.8	<0.001
Albumin (g/dL), mean±SD	3.66±0.65	3.76±0.41	3.25±0.78	3.48±0.63	0.008
SBP (mmHg), mean±SD	133.0±23.1	126.3±16.4	133.0±21.7	143.5±19.0	0.001

Variable	CKD G2 (n=40)	CKD G3 (n=40)	CKD G4 (n=40)	CKD G5 (n=40)	p value
FGF23 (pg/mL), median [IQR]	52.2 [29.9–345.9]	66.3 [29.2–121.2]	129.2 [58.7–495.6]	190.9 [103.5–695.6]	<0.001
FGF23 >1000 pg/mL, n (%)*	6 (15)	4 (10)	16 (40)	14 (35)	<0.001
log FGF23, median [IQR]	3.95 [3.40–5.84]	4.20 [3.37–4.80]	4.86 [4.07–6.19]	5.25 [4.64–6.54]	<0.001
Soluble Klotho (pg/mL), median [IQR]	48.1 [34.2–68.2]	49.5 [28.8–86.9]	55.6 [34.5–82.1]	58.2 [33.6–86.1]	0.907
Klotho <33 pg/mL (severe deficiency), n (%)	9 (23)	13 (33)	9 (23)	10 (25)	0.652
Klotho <50 pg/mL, n (%)	20 (50)	20 (50)	16 (40)	18 (45)	0.726
FGF23:Klotho ratio, median [IQR]	1.75 [0.92–3.98]	1.35 [0.68–2.28]	3.07 [1.24–7.02]	4.47 [1.56–12.53]	<0.001
CIMT (mm), mean±SD	0.898±0.087	1.004±0.084	1.164±0.107	1.283±0.116	<0.001
CIMT >1.2 mm (significant), n (%)	0 (0)	0 (0)	12 (30)	27 (68)	<0.001
ABPI, mean±SD	1.123±0.148	0.970±0.160	0.955±0.180	1.065±0.180	<0.001
ABPI <0.90 (PAD), n (%)	1 (3)	14 (35)	19 (48)	9 (23)	<0.001
LVH, n (%)	26 (65)	25 (63)	29 (73)	31 (78)	0.444
AV sclerosis, n (%)	23 (58)	26 (65)	28 (70)	15 (38)	0.018

*Abbreviations: eGFR, estimated GFR; FGF23, fibroblast growth factor 23; IQR, interquartile range; CIMT, carotid intima-media thickness; ABPI, ankle-brachial pressure index; LVH, left ventricular hypertrophy; AV, aortic valve; PAD, peripheral arterial disease. *Represents assay ceiling; actual values exceed 1000 pg/mL. p values from Kruskal-Wallis test (continuous) or chi-square test (categorical).*

FGF23 and Klotho across CKD stages: divergent trajectories
As kidney function declined, serum FGF23 rose significantly and progressively (median G2: 52.2 vs. G5: 190.9 pg/mL; H=22.09, p<0.001). In stages G2–G5, 6 (15%), 4 (10%), 16 (40%), and 14 (35%) patients reached the assay ceiling of 1000 pg/mL. This suggests that the true FGF23 gradient is much steeper than what is reflected in the measured values. This is a right-censoring phenomenon which, if anything, underestimates the difference between stages.

Across CKD stages, serum Klotho levels showed no significant difference across CKD stages (median G2: 48.1, G3: 49.5, G4: 55.6, G5: 58.2 pg/mL; H=0.55, p=0.907). Critically, 22.5–32.5% of patients in every CKD stage had severe Klotho deficiency (<33 pg/mL), and 40–50% had Klotho levels below 50 pg/mL. This flat, uniformly low Klotho trajectory across stages indicates that Klotho depletion is an early event in CKD (occurring at or before stage G2 in most patients). Thus, it cannot be expected to show a further graded cross-sectional decline in a cohort that is already selected by having CKD [4]. This is not a null finding but a mechanistically informative one: it confirms that even patients at the earliest stage in this cohort had substantially depleted Klotho, and therefore that Klotho deficiency is a near-universal, stage-independent background condition in hospitalised CKD patients.

The FGF23: Klotho ratio, which captures the degree to which FGF23 excess outpaces the Klotho buffer, showed the expected significant increase across stages (G2: median 1.75 [IQR 0.92–3.98] vs G5: 4.47 [1.56–12.53]; H=17.66, p<0.001). This divergence, rising FGF23 against a flat Klotho denominator, is the molecular basis of progressive axis uncoupling in CKD and constitutes the primary rationale for examining the FK ratio as the principal predictor variable in the analyses that follow.

Klotho tertile stratification: does Klotho modulate FGF23 cardiovascular effects?
The patients classified by Klotho tertile (<37.6, 37.6-69.4, >69.4 pg/mL) were compared on cardiovascular and biochemical parameters (Table 2). There were no significant differences found in eGFR, CKD stages, CIMT, ABPI, LVH and AV sclerosis between Klotho tertile groups. (p values: 0.188, 0.752, 0.352, 0.289, 0.960 and 0.678 respectively). As anticipated, the Klotho level alone does not independently discriminate cardiovascular status. Data already published confirm that Klotho is clinically meaningful when there is concomitant elevation of FGF23, rather than being an independent predictor.

Nevertheless, FGF23 levels were significantly greater in the low Klotho tertile (median 60.4 vs 147.2 pg/mL; p<0.001) and also the FK ratio was greater in the low Klotho group (median 2.51 vs 1.49; p=0.018). The presence of high FGF23 and low Klotho occurring in the same patients provides the biological substrate for off-target signalling.

The stratified correlation analysis reveals the crucial finding of the study. The results of the Spearman correlation analysis demonstrate a strong association between log-FGF23 and CIMT in individuals with low Klotho levels, while the association weakens at mid and high Klotho levels (r=0.642 (p<0.001) in the low-Klotho tertile; r=0.317 (p=0.021) in the middle tertile; and r=0.047 (p=0.739, non-significant) in the high-Klotho tertile) (Table 2). The step-wise attenuation – a strong correlation between FGF23 and carotid atherosclerosis that completely disappears when Klotho is present – is direct

clinical evidence of a Klotho-dependent gating effect on FGF23 vascular toxicity. Likewise, the log-FGF23 composite CV remodeling correlate may provide valuable insight regarding Klotho levels. Specifically, the correlation between log-FGF23 and the composite CV remodeling index was significant only in the low-Klotho stratum (r=0.320, p=0.018) and non-significant in both middle and high tertiles.

Table 2: Cardiovascular, biochemical, and correlation characteristics stratified by Klotho tertile.

Variable	Klotho low (<37.6 pg/mL)	Klotho moderate (37.6–69.4)	Klotho high (>69.4 pg/mL)	p value
Klotho range (pg/mL)	<37.6	37.6–69.4	>69.4	—
n	54	53	53	—
CKD stage G4–G5, n (%)	29 (54)	26 (49)	25 (47)	0.752
eGFR (mL/min/1.73m²), mean±SD	35.7±24.1	40.6±26.6	31.2±23.3	0.188
FGF23 (pg/mL), median [IQR]	60.4 [28.2–209.4]	123.5 [46.3–513.5]	147.2 [47.5–695.6]	<0.001
log FGF23, mean±SD	4.34±1.10	4.90±1.18	5.37±1.35	<0.001
FGF23:Klotho ratio, median [IQR]	2.51 [0.86–7.86]	2.21 [0.85–6.67]	1.49 [0.73–3.87]	0.018
SBP (mmHg), mean±SD	138.3±20.2	135.9±21.3	127.6±20.3	0.020
BUN (mg/dL), mean±SD	26.1±14.1	31.0±22.6	35.8±20.6	0.048
CIMT (mm), mean±SD	1.115±0.179	1.061±0.191	1.085±0.156	0.352
ABPI, mean±SD	1.017±0.192	1.054±0.158	1.014±0.182	0.289
LVH, n (%)	38 (70)	36 (68)	37 (70)	0.960
AV sclerosis, n (%)	33 (61)	31 (59)	28 (53)	0.678
r(logFGF23 vs CIMT)	**0.642*** **	**0.317***	**0.047 (ns)**	—
r(logFGF23 vs CV index)	**0.320***	**0.159 (ns)**	**0.112 (ns)**	—

*Values are mean±SD, median [IQR], or n (%) as appropriate. Spearman correlations computed within each Klotho stratum. *p<0.05; ***p<0.001; ns, not significant. CV index = z-score(CIMT) + z-score(1−ABPI) + LVH(0/1) + AV sclerosis(0/1).*

FGF23: Klotho ratio and cardiovascular remodelling

Patients were stratified according to the FK ratio tertile (T1: <1.25; T2: 1.25–4.17; T3: >4.17) to assess vascular consequences of progressive axis uncoupling (Table 3). The CIMT value in FK ratio tertiles were as follows: 1.031±0.163 mm (T1) , 1.066±0.159 mm (T2) , 1.165±0.186 mm (T3) and increased significantly across tertiles (H=16.21, p<0.001). Clinically meaningful thresholds followed a similar gradient: 52% of T1 patients had CIMT >1.0 mm, rising to 60% in T2 and 83% in T3. CIMT >1.2 mm — signifying significant atherosclerosis — was present in only 7% of T1 patients, compared to 21% of T2 and 45% of T3 patients.

In contrast, ABPI (p=0.712), LVH prevalence (p=0.903), and AV sclerosis (p=0.950) did not differ significantly across FK ratio tertiles. The selective association of axis uncoupling with CIMT but not with LVH or AV sclerosis merits discussion: CIMT captures cumulative intimal and medial vascular injury - integrating the effects of endothelial dysfunction, oxidative stress, and smooth muscle cell calcification over years - and may therefore be a more sensitive detector of long-term FGF23-mediated vascular toxicity than point-in-time binary echocardiographic assessments.

Table 3: CIMT, ABPI, LVH, and AV sclerosis across FGF23:Klotho ratio tertiles (n=160).

Variable	FK ratio T1 (low)	FK ratio T2 (moderate)	FK ratio T3 (high)	p value	Posthoc
n	54	53	53	—	—
FK ratio range	<1.25	1.25–4.17	>4.17	—	—
FGF23 median (pg/mL)	35.3	126.1	695.6	<0.001	—
Klotho median (pg/mL)	71.8	48.1	36.8	<0.001	—
CKD stage G4–G5, n (%)	16 (30)	30 (57)	34 (64)	<0.001	—
CIMT (mm), mean±SD	1.031±0.163	1.066±0.159	1.165±0.186	<0.001	T1 vs T3*
CIMT >1.0 mm, n (%)	28 (52)	32 (60)	44 (83)	0.001	—
CIMT >1.2 mm, n (%)	4 (7)	11 (21)	24 (45)	<0.001	—
ABPI, mean±SD	1.040±0.167	1.023±0.165	1.021±0.203	0.712	ns
LVH, n (%)	37 (69)	36 (68)	38 (72)	0.903	ns
AV sclerosis, n (%)	32 (59)	30 (57)	30 (57)	0.950	ns

*T1, T2, T3: tertile 1 (low), 2 (moderate), 3 (high) of the FGF23:Klotho ratio. p values from Kruskal-Wallis test (continuous) or chi-square test (categorical). *Significant post-hoc difference (Dunn's test) between T1 and T3 for CIMT.*

Four-quadrant interaction analysis: dual dysregulation and carotid remodelling

To establish the independent and interactive contributions of FGF23 excess and Klotho deficiency, we classified patients into four interaction groups according to FGF23 (upper vs lower two tertiles) and Klotho (lower vs upper two tertiles): (1) low FGF23 + high Klotho (reference, n=59); (2) high FGF23 + high Klotho (isolated FGF23 excess, n=47); (3) low FGF23 + low Klotho (isolated Klotho deficiency, n=47); and (4) high FGF23 + low Klotho (dual, n=7). The results are presented in Table 4.

Table 4: Cardiovascular remodelling outcomes by FGF23/Klotho interaction group.

Variable	LoFGF23 + HiKl (n=59) Reference	HiFGF23 + HiKl (n=47) FGF23 excess	LoFGF23 + Lo-Kl (n=47) Klotho deficient	HiFGF23 + LoKl (n=7) Dual dysregulation	p value
n	59	47	47	7	<0.001
CKD stage G4–G5, n (%)	17 (29)	30 (64)	22 (47)	7 (100)	<0.001
FGF23 median (pg/mL)	52.2	78.6	495.0	758.0	<0.001
Klotho median (pg/mL)	89.2	27.8	89.2	23.0	<0.001
FK ratio median	1.1	3.8	4.2	53.0	<0.001
CIMT (mm), mean±SD	1.065±0.181	1.078±0.162	1.083±0.170	1.364±0.069	<0.001
CIMT >1.2 mm, n (%)	5 (8)	10 (21)	17 (36)	7 (100)	<0.001
LVH, n (%)	38 (64)	35 (74)	35 (74)	3 (43)	0.218
AV sclerosis, n (%)	35 (59)	31 (66)	24 (51)	2 (29)	0.127
ABPI, mean±SD	1.045±0.170	1.016±0.189	1.020±0.176	1.026±0.239	0.739

Interaction groups defined by high/low FGF23 (upper vs lower two tertiles; threshold 211.4 pg/mL) and high/low Klotho (lower vs upper two tertiles; threshold 37.6 pg/mL). Reference group = low FGF23 + high Klotho. Hi, high; Lo, low; Kl, Klotho. CIMT p=0.0003 for dual-dysregulation vs reference (Mann-Whitney U test).

Participants in the dual-dysregulation group (high FGF23 + low Klotho) had a CIMT of 1.364±0.069 mm, which was significantly higher than that of the reference group (1.065±0.181 mm; Mann-Whitney U p=0.0003). This value was also greater than that of the high-FGF23 (1.083±0.170 mm) and low-Klotho (1.078±0.162 mm) groups. Of the dual-dysregulation group, all patients had a CIMT >1.2 mm, which was present in 8% of reference group. The average FGF23: Klotho ratio in this group was 53.0 (range ~7-125), and six of seven participants were at CKD stage G5.

Importantly, isolated FGF23 excess (high FGF23 + high Klotho) as well as isolated Klotho deficiency (low FGF23 + low Klotho) did not associate with significant differences in CIMT (mean 1.083 and 1.078 mm, respectively, p>0.05 for both vs reference). It suggests that it is the combination of both conditions and not either alone, that drives the highest vascular remodelling signal. Such a pattern supports the molecular hypothesis that Klotho normally serves to buffer vascular effects of FGF23 such that only with concurrency of FGF23 excess and Klotho depletion will off-target cardiovascular effects become clinically dominant.

The prevalence of left ventricular hypertrophy (43%) and aortic valve sclerosis (29%) was lower in the dual-dysregulation group than in the reference group (64% and 59%). However, the differences were not statistically significant. The low prevalences in the dual-dysregulation group may be due to the small sample size in this group (n=7) and advanced CKD stage in these patients, in which competing pathological processes and higher mortality risk changes phenotype.

Klotho and FGF23 univariable associations

The levels of Klotho, surprisingly, are positively correlated with FGF23 ($r = 0.371$, $p < 0.001$) and negatively with the FK ratio ($r = -0.233$, $p = 0.003$). Klotho did not correlate with eGFR ($r=-0.047$, $p=0.553$), further confirming the decoupling of Klotho from CKD stage in this cohort of patients who already have established CKD at enrolment. Positive FGF23-Klotho correlation is a known phenomenon since in CKD residual tubular Klotho expression is a target of FGF23 and high FGF23 may secondarily upregulate soluble Klotho shedding although this response is compensatory and does not maintain Klotho at sufficiently high levels to confine FGF23 to its canonical receptors [11]. Log-FGF23 correlated negatively with eGFR ($r=-0.335$, $p<0.001$), confirming relationship between declining nephron mass and rising phosphaturic hormone secretion.

Discussion

The main finding of the study is that Klotho deficiency is an important permissive factor that enhances the carotid vascular effects of elevated FGF23 in chronic kidney disease (CKD). In patients with low Klotho, log-FGF23 and CIMT were strongly correlated ($r=0.642$, $p<0.001$), whereas, in those with an intermediate Klotho, the correlation was moderate ($r=0.317$, $p=0.021$). Notably, in patients with a replete Klotho, the correlation was completely non-significant ($r=0.047$, $p=0.739$). This Klotho-stratified, gradient-dependent attenuation of the FGF23-vascular association is the direct clinical correlate of the molecular gating hypothesis first proposed in experimental models. [11]

A well-established mechanistic basis for this observation exists in the experimental literature. The soluble klotho restrains the sodium-dependent phosphate co-transporters (NaPi) found in VSMCs, preventing the phosphate-induced osteogenic trans differentiation and mineralisation [5]. It also promotes the activity of eNOS and inhibits Wnt/β-catenin and TGF-β signalling pathways, thereby preserving the integrity of the endothelium and countering the pro-fibrotic and pro-inflammatory environment that accelerates intima-media thickening [9,12]. Klotho depletion leads to the simultaneous withdrawal of these protective mechanisms. Elevated FGF23, which does not have renal co-receptor, then gains access to FGFR1 and FGFR4 on endothelial and smooth muscle cells in a Klotho-independent manner and activates pro-inflammatory pathways including nuclear factor kappa-B (NF-κB) and raises reactive oxygen species generation [12,13]. Cumulative burden of endothelial dysfunction, smooth muscle proliferation and medial calcification would make CIMT as the appropriate clinical readout for this prolonged, multi-mechanism vascular injury.

The fact that neither isolated FGF23 excess nor isolated Klotho deficiency were associated with CIMT elevation, yet those of dual-dysregulation had associated markedly elevated CIMT (1.364 mm, all >1.2 mm) making it a significant finding. This pattern of interaction constitutes the molecular prediction: FGF23 is vasculotoxic when Klotho is insufficient to restrain it to the canonical receptors and Klotho deficiency is fully detrimental when there is sufficient FGF23 to activate off-target pathways. It should be noted that this finding ought to be taken with caution due to the small size of the dual-dysregulation group (n=7), constrained by the joint probability of extreme values on both biomarkers. This was present in a 160-patient cohort. Larger cohorts are needed to verify the interaction.

The selective association of CIMT, and not LVH, AV sclerosis or ABPI, with axis-related associations requires explanation. There are many considerations. To begin, the occurrence of CIMT is a continuous and high-resolution vascular measure that builds up a pathological signal over years; therefore, it is a more sensitive detector of long-term endocrine toxicity compared to point-in-time echocardiographic binary measurements. Moreover, LVH is already prevelant in 65% of G2 patients in our cohort, due to high background rates of hypertension and diabetes, thereby creating a ceiling effect. Thus, no additional FGF23-FGFR4-calcineurin-NFAT contribution is detectable. Third, the ABPI displayed a non-monotonic pattern consistent with medial calcification

in advanced CKD. This effect may paradoxically increase ABPI in patients with G5 and highest FK ratios, blunting the anticipated negative correlation [14].

The clinical interpretation of the research findings is strongly influenced by the flat Klotho trajectory across CKD stages, which has been noted in different studies [4,11]. The absolute amount of Klotho cannot be used for CKD staging or individual risk prediction due to the necessity of FGF23 measurement. As a result, the relative axis imbalance reflected in the FK ratio is clinically more informative than Klotho or FGF23 measured in isolation. Based on our data, both CKD stages and CIMT tertile demonstrated a notable clear gradient of the FK ratio and its tertile classification cleanly stratified the proportion of patients with CIMT >1.2 mm (7% vs 21% vs 45% T1, T2, T3). The FK ratio should be an addition to the CKD-cardiovascular risk model, as a composite biomarker.

Once we consider the pathophysiological context, the unexpected positive association between klotho and FGF23 ($r=0.371$, $p<0.001$) in this cohort is no longer paradoxical. In chronic kidney disease (CKD), FGF23 is released in response to the combined impact of phosphate load, PTH, and inflammatory cytokines. Simultaneously, FGF23 binding to residual tubular klotho and the resultant downstream signalling may quickly upregulate the shedding of soluble klotho from the tubular surface. Moreover, the higher FGF23 levels in patients of this cohort tended towards more advanced CKD stages (where the distributions of FGF23 and Klotho overlap widely), leading to a false positive association in cross-sectional data. One should not interpret this as protective of higher Klotho against the toxicity of FGF23, it just reveals how the axis is regulated in concert, rather than any ceiling effect.

This study faced a number of limitations. The cross-sectional design allows correlation but not causal inference; we cannot establish that Klotho deficiency temporally precedes or causes the vascular changes nor that FGF23 elevation directly mediates them rather than being a common correlate of CKD severity. With only seven participants, the dual-dysregulation group is too small to draw strong conclusions; future studies should seek to prospectively identify and enrich this high-risk phenotype. All the cardiovascular measures include the CIMT, ABPI, LVH, AV sclerosis which are structural rather than functional and may not accurately assess the haemodynamic consequences of off-target FGF23 signalling. The research took place in a single tertiary-care centre, predominantly housing a hospitalized population of South Asian origin, and the absolute biomarker thresholds may not apply to other populations. Lastly, the standardisation of assays for measuring Klotho remains a challenge, and the ELISA values of the present study should be interpreted in the light of the assay itself.

Regardless of these limitations, this study provides what, to our knowledge, is the first systematic clinical evidence for Klotho-dependent gating effect on FGF23 cardiovascular toxicity in a CKD cohort. The stratified correlational approach used here, showing a strong FGF23-CIMT association in Klotho-depleted patients which completely disappears in Klotho-replete patients, translates the molecular gating hypothesis directly to humans, without requiring biomarker manipulation or animal modelling.

This is significant in therapy. Strategies that are currently being employed in order to target the excess FGF23 in CKD include various dietary phosphate restrictions, administering phosphate binders, the use of FGF23 antibodies such as burosumab, etc. However, if Klotho remains depleted, these strategies may fail. This is because the downstream off-target receptors remain accessible in such cases. In contrast to the aforementioned strategies, giving recombinant soluble Klotho (currently in phase I clinical trials) or taking pharmacological agents such as angiotensin-converting enzyme inhibitors, PPAR-γ agonists or rapamycin [15] may also protect against FGF23 cardiovascular toxicity without lowering FGF23 itself by restoring receptor gating. The clinical finding that the FGF23-CIMT correlation is essentially zero in patients with replete Klotho supports this strategy: if Klotho can be restored to the high-tertile range, it is plausible that the carotid atherosclerotic signal from FGF23 elevation will be substantially abrogated.

Conclusions

This clinical and molecular correlation study provides evidence that Klotho deficiency is a critical permissive factor in the off-target cardiovascular toxicity of elevated FGF23 in CKD. The strength of the FGF23–carotid IMT association is Klotho-dependent, grading from a strong positive correlation ($r=0.642$) in Klotho-depleted patients to complete non-significance ($r=0.047$) in Klotho-replete patients. The FGF23:Klotho ratio, capturing the degree of axis uncoupling, is more informative than either marker alone as a correlate of vascular remodelling and increases significantly with advancing CKD stage. Dual FGF23/Klotho dysregulation identifies a small but extreme-risk subgroup with severe carotid remodelling. These findings support the concept that Klotho restoration - rather than FGF23 suppression alone - is the therapeutically rational approach to attenuating off-target cardiovascular effects in CKD. External prospective validation with longitudinal cardiovascular outcome data is the essential next step.

Declarations

Conflict of interest

The authors declare no conflicts of interest.

Funding

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

Ethical approval

Institutional ethics committee approval was obtained IEC No. CSP-MED/24/JAN/97/19. All participants provided written

informed consent.

Data availability

The anonymised dataset is available from the corresponding author on reasonable request.

AI usage

Data was analysed using Claude AI and further verified by a statistician.

Role of authors

Dr. Swathy Moorthy - Conceptualization, Methodology, Data Curation, Formal Analysis, Resources, Writing – Original Draft Preparation, Review & Editing, Visualization, Supervision, Project Administration.

Dr. Subha Sree S - Conceptualization, Methodology, Data Curation, Formal Analysis, Resources, Writing – Original Draft Preparation, Review & Editing, Visualization.

References

1. Urakawa I, Yamazaki Y, Shimada T, Iijima K, Hasegawa H, Okawa K, Fujita T, Fukumoto S, Yamashita T. Klotho converts canonical FGF receptor into a specific receptor for FGF23. Nature. 2006;444(7120):770–774.
2. Komaba H, Fukagawa M. The role of FGF23 in CKD: with or without Klotho. Nat Rev Nephrol. 2012;8(9):484–490.
3. Isakova T, Wahl P, Vargas GS, Gutiérrez OM, Scialla J, Xie H, Appleby D, Nessel L, Bellovich K, Chen J, Hamm L, Gadegbeku C, Horwitz E, Townsend RR, Anderson CAM, Lash JP, Hsu CY, Leonard MB, Wolf M; CRIC Study Group. Fibroblast growth factor 23 is elevated before parathyroid hormone and phosphate in chronic kidney disease. Kidney Int. 2011;79(12):1370–1378.
4. Hu MC, Shi M, Zhang J, Quinones H, Griffith C, Kuro-O M, Moe OW. Klotho deficiency causes vascular calcification in chronic kidney disease. J Am Soc Nephrol. 2011;22(1):124–136.
5. Lim K, Lu TS, Molostvov G, Lee C, Lam FT, Zehnder D, Hsiao LL. Vascular Klotho deficiency potentiates the development of human artery calcification and mediates resistance to fibroblast growth factor 23. Circulation. 2012;125(18):2243–2255.
6. Faul C, Amaral AP, Oskouei B, Hu MC, Sloan A, Isakova T, Gutiérrez OM, Aguillon-Prada R, Lincoln J, Hare JM, Mundel P, Morales A, Scialla J, Fischer M, Soliman EZ, Chen J, Go AS, Rosas SE, Nessel L, Townsend RR, Feldman HI, St John Sutton M, Ojo A, Gadegbeku C, Di Marco GS, Reuter S, Kentrup D, Tiemann K, Brand M, Hill JA, Moe OW, Kuro-o M, Kusek JW, Keane MG, Wolf M. FGF23 induces left ventricular hypertrophy. J Clin Invest. 2011;121(11):4393–4408.
7. Grabner A, Amaral AP, Schramm K, Singh S, Sloan A, Yanucil C, Li J, Shehadeh LA, Hare JM, David V, Martin A, Fornoni A, Di Marco GS, Kentrup D, Reuter S, Mayer AB, Pavenstädt H, Stypmann J, Kuhn C, Hille S, Frey N, Leifheit-Nestler M, Richter B, Haffner D, Abraham R, Bange J, Sperl B, Ullrich A, Brand M, Wolf M, Faul C. Activation of cardiac fibroblast growth factor receptor 4 causes left ventricular hypertrophy. Cell Metab. 2015;22(6):1020–1032.
8. Han X, Quarles LD. FGF23 induced left ventricular hypertrophy mediated by FGFR4 signaling in the myocardium is attenuated by soluble Klotho in mice. J Mol Cell Cardiol. 2020;138:67–80.
9. Tan SJ, Smith ER, Hewitson TD, Holt SG, Toussaint ND. The importance of klotho in CKD and cardiovascular disease: a review. Clin Kidney J. 2018;11(6):731–742.
10. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, Crews DC, Doria A, Estrella MM, Froissart M, Grams ME, Greene T, Grubb A, Gudnason V, Gutiérrez OM, Kalil R, Karger AB, Mauer M, Navis G, Nelson RG, Poggio ED, Rodby R, Rossing P, Rule AD, Selvin E, Seegmiller JC, Shlipak MG, Torres VE, Yang W, Ballew SH, Couture SJ, Powe NR, Levey AS; CKD-EPI Collaboration. New creatinine- and cystatin C–based equations to estimate GFR without race. N Engl J Med. 2021;385(19):1737–1749.
11. Kuro-o M, Matsumura Y, Aizawa H, Kawaguchi H, Suga T, Utsugi T, Ohyama Y, Kurabayashi M, Kaname T, Kume E, Iwasaki H, Iida A, Shiraki-Iida T, Nishikawa S, Nagai R, Nabeshima YI. Mutation of the mouse klotho gene leads to a syndrome resembling ageing. Nature. 1997;390(6655):45–51.
12. Vervloet MG, Olgaard K, Lewin E, Heijboer AC, van Zuilen AD. Direct, acute effects of Klotho and FGF23 on vascular smooth muscle and endothelium. PLoS One. 2014;9(4):e93423.
13. Leifheit-Nestler M, Grabner A, Hermann L, Richter B, Schmitz K, Fischer DC, Yanucil C, Faul C, Haffner D. Induction of cardiac FGF23/FGFR4 expression is associated with left ventricular hypertrophy in patients with chronic kidney disease. Nephrol Dial Transplant. 2016;31(7):1088–1099.
14. Aboyans V, Ricco JB, Bartelink MEL, Björck M, Brodmann M, Cohnert T, Collet JP, Czerny M, De Carlo M, Debus S, Espinola-Klein C, Kahan T, Kownator S, Mazzolai L, Naylor AR, Rees O, Röther J, Sprynger M, Tendera M, Tepe G, Venermo M, Vlachopoulos C, Desormais I; ESC Scientific Document Group. 2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases. Eur Heart J. 2018;39(9):763–816.

15. Grabner A, Schramm K, Silswal N, Hendrix M, Yanucil C, Czaya B, Singh S, Wolf M, Hermann S, Stypmann J, Di Marco GS, Brand M, Wacker MJ, Faul C. FGF23/FGFR4-mediated left ventricular hypertrophy is reversible. Sci Rep. 2017;7(1):1993.

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

Longitudinal Laboratory and Diagnostic Profile of Patients with Prostate Cancer and PCA3 Expression: A Case Series

Divye Ahlawat[1], Aditya Khatri[1], Amanpreet Singh Gulian[2], Debprosit Karmakar[1], Sakshi Rai[1], Shubhangi Mukherjee[1], Suman Kumar Ray[1], Ketan Mehra[2], Sukhes Mukherjee[1*]

[1]*Department of Biochemistry. All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India*
[2]*Department of Urology. All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India*

Article Info

[*]***Corresponding Author:***
Sukhes Mukherjee
Department of Biochemistry.
All India Institute of Medical Sciences, Bhopal, Madhya Pradesh-462020, India
E-mail: sukhes.biochemistry@aiimsbhopal.edu.in

Keywords

Prostate cancer, PCA3, Gleason scores, biomarkers, disease progression and monitoring

Abstract

Background: Prostate cancer remains one of the most prevalent malignancies affecting men worldwide, underscoring critical need for robust monitoring and diagnostic procedures. This study investigates dynamic biological landscape of the prostate cancer and its systemic effects by longitudinally evaluating the hematological, hepatic, and renal profiles of the 17 patients compared against 21 control samples.

Methods: A retrospective review was conducted on the 17 patients with histologically verified prostate cancer and 21 control samples. The laboratory evaluations included hematology, prostate-specific markers (PCA3, Gleason scores), inflammatory markers (CRP), liver and renal function tests. The data from multiple visits enabled trend and correlation analyses while statistical analysis determined marker-tumor relationships and variances between patient and control groups.

Results: The PCA3 expression levels in the patients varied between 0.1 and 0.6, while tumor samples exhibited high-grade characteristics with the Gleason scores between 6 and 10 (primarily centered on 8). The Compared to 21 control samples, patients' hemoglobin levels (13.5–14.2 g/dL) remained within normal to the slightly anemic ranges, indicating a satisfactory systemic health status. Patient CRP levels varied between 12.0 and 40.0 mg/L, suggesting sporadic inflammatory activity not observed in the controls. Hepatic (bilirubin, AST, ALT, ALP) and renal (urea, creatinine) markers in the patients showed only minor variations and remained within normal parameters, comparable to control cohort. PCA3 levels correlated with the tumor aggressiveness but showed no significant relationship with other systemic tests.

Conclusion: This longitudinal data implies that baseline systemic health condition of the prostate cancer patients—as defined by hematological, hepatic, and renal parameters—remains stable and comparable to 21 control samples. However, intermittent fluctuations in the CRP and other inflammatory markers suggest a distinct systemic response tied to the tumor activity. Ultimately, combining these

standard clinical parameters with molecular and inflammatory markers may significantly improve the targeted management of prostate cancer.

Introduction

Prostate cancer is one of the most common and difficult forms of cancer men experience worldwide. It is predicted that by 2040, the PC burden will have increased to approximately 2.4 million cases and 712 000 deaths solely because of the aging and growing global population [1]. Its complex biology, variable clinical presentation, and diverse progression pattern warrant comprehensive diagnostic, prognostic, and monitoring strategies to ensure better patient outcomes. With the progression of time and advances in molecular diagnostics, traditional biochemical, and histopathological tests, we have been able to understand the disease process and develop more individualized approaches to manage it [2]. Yet, with the progression of prostate cancer, the heterogeneity of the disease process remains one of the significant challenges, emphasizing the need to employ a variety of biomarkers and laboratory tests [3]. This study aims to assess the longitudinal data on prostate cancer patients and their respective hematological, liver, kidney, and tumor markers like Gleason score. Prostate Cancer Antigen 3 (PCA3) is a non-coding RNA, which is overexpressed in prostate cancer cells [4]. It acts as a major biomarker for the diagnosis and monitoring of the disease, especially when the level of PSA is undetermined. PCA3 is a non-invasive diagnostic test conducted on urine samples collected after Digital Rectal Examination, and its specificity is higher than the level of PSA for differentiating malignant and benign prostate diseases [5]. On the other hand, the Gleason score is a histopathological grading scale used to measure the severity of prostate cancer [6]. This scale varies from 6 to 10. Higher values indicate aggressive and undifferentiated forms of prostate cancer. The integration of the PCA3 gene and the Gleason score will help in the accurate diagnosis and prognosis of prostate cancer [7]. This will, in turn, help in the treatment process for patients suffering from prostate cancer. The relationship between the PCA3 gene and the Gleason score is established on the basis of the fact that the severity of the disease is assessed using the combined effect of the two. It has been observed that the values obtained for the PCA3 gene in the patients increase the values obtained for the Gleason score. This establishes the fact that the PCA3 gene has the potential to act as a non-invasive tool for the detection of prostate cancer, especially the high-grade type. The integration of the expression of the PCA3 gene and the Gleason score will help in the accurate diagnosis and prognosis process, which will, in turn, help in the treatment process for the patient.

The hematological parameters of hemoglobin, RBC count, hematocrit, WBC count, and platelets were also found to be within the normal range or slightly low, indicating the general health of the body [8]. This is a positive sign, implying that condition may be more responsive to the traditional therapy procedures without added danger of the hematological problems. The hyperactive immune responses and persistent tissue irritation are the significant contributors in evolution of the cancer. High levels of the C-reactive protein (CRP) interacted favorably with the PCA3 and Gleason scores, emphasizing relationship between the systemic cellular stress and prostate cancer progression [9, 10]. This also supports concept that immune-regulated microenvironments play a significant role in development of the prostate cancer [11]. HbA1c levels, which indicate the degree of glycaemic control, indicates that the glycaemia is one of the controlling factors in the diabetic patients. Although the significance of the metabolic syndrome is becoming increasingly recognized with regards to the risk of prostate cancer as well as the prognosis of the disease [12, 13] the glycaemic control was well-managed in the patients of the present study cohort, thus minimizing the confounding effect of hyperglycemia or insulin resistance. These biochemical profiles thus emphasize the diversity of prostate cancer. Although the Gleason score is high, the fact that the hematological, hepatic, as well as renal functions of the patients are normal suggests that the patients are still quite stable, thus giving the doctor a chance to treat the patients [14]. Relationship between the systemic immune activity and tumor indicators such as PCA3 and Gleason score highlights importance of these hematological biomarkers in disease prognosis [15-17].

Materials and methods

Ethical clearance

Ethical clearance was obtained from the Institute Ethics Committee (AIIMSBPL/IHECSR/Nov-2024/29 dated: 23rd Dec 2024) of AIIMS Bhopal, India. and all subjects gave their informed consent before participation. From the Department of Urology, AIIMS Bhopal, patients were recruited based on the initial checkup and to participate in the study, after fulfilling the inclusion and exclusion criteria and after obtaining consent. All studies involving human subjects must indicate that they are in compliance with the ethical principles for medical research involving human subjects, in accordance with the Declaration of Helsinki.

Study design and population

This study was conducted as a retrospective case series with 17 male patients diagnosed with the prostate cancer and a comparison cohort of 21 control samples. Research population was the selected from AIIMS Bhopal's medical records database (figure 1), which contained the information from January to June 2025. The inclusion criteria included male patients aged 50 and older who had a confirmed histopathological diagnosis of the prostate cancer and had also undergone routine laboratory evaluations as part of their ongoing clinical management, while 21 controls were matched based on similar demographic profiles and the baseline lab data. Patients with the co-existing cancers, significant

medical conditions that may influence lab values (such as active infections, autoimmune conditions, or hematological conditions), insufficient medical records were excluded to the ensure data accuracy and reliability and to keep the focus on prostate cancer-related changes.

Figure 1: Location where the study is conducted.



The major aim was to examine the data from the patients' laboratory records over time and determine if there were any associations between prostate-specific markers, the patients' inflammatory reactions, liver function, kidney function, and blood parameters. Ethical considerations in the study of human subjects were strictly adhered to, with approval from the ethics committee prior to the collection of the data.

Data collection

The data was gathered by conducting a thorough evaluation of electronic health records of 17 the patients and 21 control samples utilizing hospital's health information system. This study collected laboratory data longitudinally at the several time periods, such as routine checks, diagnostic testing intervals and follow-up visits. The patient cohort's data was grouped chronologically based on time of diagnosis, therapy, illness development and it was compared to baseline values set by the 21 control samples.

This study involved a wide range of hematological, biochemical, and tumor-specific markers. In the hematological tests, the quantification of hemoglobin, red blood cell (RBC) counts, hematocrit, white blood cell (WBC) counts, and platelets was done. The study also involved the quantification of PCA3 and the Gleason score from the biopsy reports to establish the aggressiveness and advancement of the tumors. In the study, the quantification of CRP was done to establish the level of systemic inflammation. In the study, the following were done to establish the level of activity of the liver: total bilirubin, direct bilirubin, indirect bilirubin, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, total protein, albumin, globulin, and the albumin/globulin ratio. Renal function was assessed by determining the amount of urea, creatinine, sodium, and potassium in the blood. Glycemic control was assessed by measuring the level of HbA1c.

Laboratory analysis

The laboratory parameters used in this study were rigorously analyzed following standardized clinical laboratory protocols, which ensured the accuracy of the results and their reproducibility during different sessions of the tests. Such protocols are extremely significant in ensuring that laboratory results are interpreted in an accurate and standardized fashion. The automated hematology analyzer was used to determine the RBC counts, hematocrit, WBC counts, and platelet counts. We obtained the levels of PCA3 using reports from molecular diagnostic tests with the PCA3 Detection Kit(Prostate cancer antigen 3, Code: EH4353 and Batch No: FN250827) that used molecular methods like nucleic acid amplification tests. The Gleason score, which was obtained from the histopathological examination of the biopsy sample and the score, which ranged from 6 to 10, was correlated with the other test results to know the severity of the disease. We used high-sensitivity immunoassays like enzyme-linked immunosorbent assay (ELISA) to know the level of CRP in the blood. This test used

antibodies to detect the presence of small amounts of CRP in the blood. The tests for liver functions that we conducted included total bilirubin, direct bilirubin, and indirect bilirubin, which we conducted using a spectrophotometer. These tests involve chemical reactions that change the color of the dye when they come into contact with bilirubin. The change in color is then measured using a spectrophotometer to obtain an accurate value for bilirubin. We conducted enzymatic and colorimetric tests to determine the level of liver enzymes, which included AST, ALT, and ALP. These tests, which measure the level of liver enzymes, use a reaction known as a substrate conversion reaction, which produces a signal that can be measured and is directly proportional to the enzyme. Enzymatic tests were used to measure urea and serum creatinine. In our current study, we used high-performance liquid chromatography (HPLC) to measure the level of HbA1c.

Statistical analysis

Data were presented as mean ± SD. The statistical comparisons were made to the determine differences and coefficient of the variation (CV) in laboratory parameters between 17 prostate cancer patients and 21 control samples. Biochemical indicators were examined throughout time with the statistical significance set at $p < 0.05$.

Results

Patient demographics and prostate markers

The patient population's PCA3 levels ranged from 0.1 to 0.6, where a lower score closer to 0.1 indicates less active, slower-growing cancer cells and a higher score closer to 0.6 signals more active disease. This indicates that the prostate cancer among the patients was at different stages of activity. The range of Gleason scores among the patients was between 6 and 10. This indicates that prostate cancer among the patients was at different stages of differentiation. Most of the patients had a Gleason score of 8 or more. The variation in PCA3 levels among the patients is related to the variation in Gleason scores. This indicates that prostate cancer with high levels of PCA3 is more related to poor prognosis. All these markers indicate the heterogeneity of prostate cancer among the patients. This indicates the need to evaluate prostate cancer among patients individually. All these findings indicate how complex the biology of prostate cancer is among the patients.

Hematological profile

Statistical comparisons of 17 prostate cancer patients and 21 control samples indicated unique characteristics across parameters tested. Average PCA3 level in patient group was 0.36 ±0.15 (range: 0.10-0.60), indicating high variability with a Coefficient of Variation (CV) of 42.22%. The average Gleason score was 8.1 ±1.4 (median: 8; range: 6-10), suggesting mostly high-grade tumors with the limited variability (CV: 17.28%). Patients' average the CRP level was 24.3 ±10.7 mg/L (range: 12-45 mg/L). High CV of the 44.05% indicates a very variable and the sporadic systemic inflammatory response in patient group, which was substantially higher than steady baseline of the 21 control samples ($p<0.05$). In contrast, patient hematological values remained within or near normal reference limits, with no statistically significant evidence of severe anemia or leukocytosis compared to the 21 controls ($p > 0.05$). Patients had a mean hemoglobin level of 13.9 ±0.5 g/dL (range: 13.5-14.2 g/dL) with minimal variability (CV: 3.60%). The remaining patient hematological averages demonstrated low to moderate variability: Red blood cell count averaged approximately 4.72 million/µL, the hematocrit remained constant between 42% and 45% (mean: ~43%) the white blood cell count averaged 6,900/mm³ (range: 6,000-8,000/mm³), and average platelet count was 244,000/mm³ (range: 220,000-280,000). Overall, difference in the CV values between very stable hematological markers and highly variable inflammatory responses highlights unique biological alterations associated with the tumour activity in patient group.

Table 1: Summary of parameter distributions and statistical measures.

Parameter	Patient Mean ± SD (n=17)	CV (%)	pvalue (vs. Controls, n=21)
PCA3	0.36 ±0.15	42.22	N/A
Gleason Score	8.1 ±1.4	17.28	N/A
CRP (mg/L)	24.3 ±10.7	44.05	< 0.05
Hemoglobin (g/dL)	13.9 ±0.5	3.60	> 0.05
RBC (Million/µL)	4.72 ±0.31	6.58	> 0.05
Hematocrit (%)	43.0 ±1.9	4.42	> 0.05
WBC (/mm³)	6900 ±1032	15.00	> 0.05
Platelets (/mm³)	243,529 ±23,727	9.75	> 0.05

Inflammatory marker and Gleason score

CRP values in the patient group varied from 12.0 to 45.0 mg/L, indicating sporadic systemic inflammatory activity. Molecular analysis revealed the PCA3 values ranging from the 0.1 to 0.6 (mostly clustered between 0.3 and 0.4) and indicating moderate expression levels among patients. The histopathological examination indicated high-grade illness with Gleason scores ranging from 6 to 10, with majority of the patients scoring 8 or 9. Patients with the higher Gleason scores (8 to 10) had persistently raised the CRP levels, showing a direct relationship between the advanced tumor aggressiveness and the systemic inflammation. A correlation the study revealed a significant positive connection between PCA3 expression and tumor severity as measured by the Gleason score ($r = 0.62$, $p < 0.001$). Similarly, systemic immunological activity evaluated by the CRP levels showed a substantial positive connection with the Gleason score ($r = 0.55$, $p = 0.003$). In the contrast, baseline hematological indicators, such as hemoglobin levels and the WBC counts, did not demonstrate a significant link with the Gleason scores or other tumor characteristics ($p > 0.05$), the indicating that they have no diagnostic utility for the tumor classification. Taken together our data show that, while hematological markers stay steady, PCA3 and CRP are the important, dependable predictors of the prostate cancer severity and tumor aggressiveness.

Table 2: Correlations among PCA3, Gleason Score, and CRP.

Parameter	Mean ± SD (Range)	Statistical findings	Clinical relevance
PCA3	0.36 ±0.15 (0.10–0.60)	Correlation with Gleason score: $r = 0.62$, $p < 0.001$	Significant positive association; PCA3 tracks closely with tumor aggressiveness.
CRP (mg/L)	24.3 ±10.7 (12–45)	Correlation with Gleason score: $r = 0.55$, $p = 0.003$	Significant positive association; linked to tumor severity.
Gleason Score	8.1 ±1.4 (6–10)	Strongly correlated with both PCA3 and CRP ($p < 0.01$)	Higher tumor grades are significantly associated with elevated molecular and inflammatory markers.

Liver function tests

Liver function testing on the 17 prostate cancer patients demonstrated that the liver enzymes and functional indicators remained within the healthy clinical standard ranges, indicating stability comparable to 21 control samples ($p > 0.05$) (table 3). Total bilirubin (14.8 ±4.8 mg/dL), direct bilirubin (7.7 ±2.8 mg/dL) and indirect bilirubin (7.1 ±2.9 mg/dL) showed minor but is not statistically significant changes. The liver enzymes AST (42.9 ± 7.7 U/L) and the ALP (130.7 ± 9.4 U/L) were normal. Although ALT levels (117.2 ±8.4 U/L) were within permissible limits and there was a slight, statistically discernible difference compared to the controls ($p = 0.05$). All participants had rather constant levels of the total protein (8.4 ±0.3 g/dL), albumin (4.2 ±0.1 g/dL), globulin (4.2 ± 0.2 g/dL) and the estimated albumin/globulin (A/G) ratio (1.04 ±0.04). Overall, the modest coefficients of the change across these metabolic indicators suggest that baseline hepatic function was conserved in patient group despite aggressive tumor growth.

Table 3: Summary of liver function tests and statistical comparisons.

Parameter	Patient Mean ± SD (n=17)	CV (%)	pvalue (vs. Controls, n=21)
Bilirubin Total (mg/dL)	14.8 ±4.8	32.43	> 0.05
Bilirubin Direct (mg/dL)	7.7 ±2.8	36.36	> 0.05
Bilirubin Indirect (mg/dL)	7.1 ±2.9	40.85	> 0.05
AST (U/L)	42.9 ±7.7	17.95	> 0.05
ALT (U/L)	117.2 ±8.4	7.17	≤0.05
ALP (U/L)	130.7 ±9.4	7.19	> 0.05
Total Protein (g/dL)	8.4 ±0.3	3.57	> 0.05
Albumin (g/dL)	4.2 ±0.1	2.38	> 0.05
Globulin (g/dL)	4.2 ±0.2	4.76	> 0.05
A/G Ratio	1.04 ±0.04	3.85	> 0.05

Renal function tests

The renal function testing showed fairly stable, well-preserved kidney function in all the 17 prostate cancer patients. All standard parameters remained within the normal clinical reference limits. The average serum urea level was 21.6 ±1.7 mg/dL, ranging from 19 to 23 mg/dL. The average creatinine level was 1.27 ±0.15 mg/dL, with a range of 1.0 to 1.4 mg/dL. Both had low coefficients of variation, at 7.87% and 11.81%, respectively. Statistical comparisons between the patient group and the 21 control samples found no significant differences ($p > 0.05$). Serum electrolytes also showed similar stability and balance, with an average of 137.0 ±1.0 mmol/L for sodium and 4.75 ±0.20 mmol/L for potassium. These findings indicate that baseline kidney function and electrolyte balance remained the same in the patient group without any compromise. This remained true even with high-grade tumors (mean Gleason score: 8.0 ±1.3) and a strong systemic immune response (mean CRP: 22.4 ±8.2 mg/L).

Table 4: Summary of renal function tests and statistical comparisons.

Parameter	Patient Mean ± SD (n=17)	CV (%)	pvalue (vs. Controls, n=21)
Urea (mg/dL)	21.6 ±1.7	7.87	> 0.05
Creatinine (mg/dL)	1.27 ±0.15	11.81	> 0.05
Sodium (Na, mmol/L)	137.0 ±1.0	0.73	> 0.05
Potassium (K, mmol/L)	4.75 ±0.20	4.21	> 0.05

Glycemic control

The differences between healthy men and men with prostate cancer were analyzed and connection analyses were conducted to assess the degree of connection between the markers. By contrasting a sample group (n = 17) with a control group (n = 21), our study offers an investigation of prostate cancer features in relation to inflammatory and metabolic indicators. With a mean HbA1c of 6.05% ±0.49% (CV = 8.10%) that was completely within the 5.5% to 6.4% range, the sample showed well-controlled blood sugar levels. With a mean value of 0.36 ±0.15 (CV = 41.67%) and a statistically significant, strong positive correlation with C-reactive protein (CRP) levels ($p = 0.025$), the prostate cancer marker PCA3 is further supported by elevated systemic inflammation, with a mean CRP level of 22.1 mg/L ±7.8 (CV = 35.29%) that varied significantly against the reference baseline ($p = 0.02$). Tumor aggressiveness was found to be high in the study population as shown by the mean of the Gleason score of 8.0 ±1.2, with a CV of 15.00%. While the mean of the Gleason score was significantly correlated with the serum level of CRP ($p = 0.20$), no connection was noted between HbA1c and the studied parameters that indicated systemic immune responses ($p = 0.30$). Meanwhile, the mean values and variations between the levels of the PCA3 and CRP significantly highlighted their potential and large involvement in the progression of prostate cancer.

Table 5: Clinical and molecular parameters (Sample n=17, Control n=21).

Parameter	Mean ± SD	Coefficient of Variation (CV %)	Correlation with CRP (pvalue)	Comparison vs. Reference (pvalue)
PCA3	0.36 ±0.15	41.67%	p = 0.025	—
Gleason Score	8.0 ±1.2	15.00%	p = 0.20	—
CRP (mg/L)	22.1 ±7.8	35.29%	—	p = 0.02
HbA1c (%)	6.05 ±0.49	8.10%	p = 0.30	

The data represents a population with high-grade tumor pathology and systemic inflammation compared to the baseline organ functioning. The mean Gleason Score (8.1 ± 1.4) shows high-grade disease and it is significantly ($p < 0.05$) associated with elevated C-Reactive Protein (CRP: 24.3 ± 10.7 mg/L), showing a strong, intermittent, systemic inflammatory response. PCA3 also shows large patient heterogeneity with CV: 42.22% but does correlate with inflammation and tumor grade severity. Despite elevated disease markers, these patients maintained stable homeostatic balance. Renal function tests show preserved glomerular clearance and sodium, the most tightly controlled homeostatic parameter, displays the narrowest range (CV: 0.73%). Hematological parameters also remain within normal limits with hemoglobin, red and white blood cells and platelets showing no important differences from baseline. Liver function tests evidenced maintained hepatic structural integrity and generally well-preserved synthetic function with a mild, statistically important elevation in ALT ($p <0.05$). No connection was observed between systemic inflammation in this cohort and overall glycemic control or metabolic health (mean HbA1c: 6.05 ±0.49%).

Table 6: Comprehensive clinical evaluation of patient demographics, molecular markers, hematological profiles, liver function, renal function, and glycemic control in the study cohort.

Parameter category	Specific parameter	Patient mean ± SD	Coefficient of Variation (CV %)	Statistical significance (pvalue)	Key findings
Patient demographics and molecular markers	PCA3	0.36 ±0.15	42.22%	N/A	High heterogeneity; strongly correlates with Gleason Score (r=0.62, p<0.001) and CRP (p=0.025).
	Gleason Score	8.1 ±1.4	17.28%	N/A	Predominantly highgrade tumors (range 6–10); correlates significantly with elevated CRP (r=0.55, p=0.003).
Hematological profile	CRP (mg/L)	24.3 ±10.7	44.05%	p < 0.05	High, sporadic systemic inflammatory response; significantly elevated vs. stable control baseline.
	Hemoglobin (g/dL)	13.9 ±0.5	3.60%	p > 0.05	Stable; within normal limits; no signs of severe anemia.
	RBC (Million/µL)	4.72 ±0.31	6.58%	p > 0.05	Stable; no significant variance vs. controls.
	Hematocrit (%)	43.0 ±1.9	4.42%	p > 0.05	Highly constant values across the patient cohort.
	WBC (/mm³)	6900 ±1032	15.00%	p > 0.05	Normal range; rules out active leukocytosis.
	Platelets (/mm³)	243,529 ±23,727	9.75%	p > 0.05	Normal and stable baseline values.
Liver function tests (LFT)	Bilirubin Total (mg/dL)	14.8 ±4.8	32.43%	p > 0.05	Normal hepatic clearance with mild, minor fluctuations.
	Bilirubin Direct (mg/dL)	7.7 ±2.8	36.36%	p > 0.05	Normal range; values show minor, insignificant variability.
	Bilirubin Indirect (mg/dL)	7.1 ±2.9	40.85%	p > 0.05	Values display moderate variation but remain nonsignificant.
	AST (U/L)	42.9 ±7.7	17.95%	p > 0.05	Intact hepatic structural integrity.
	ALT (U/L)	117.2 ±8.4	7.17%	p <0.05	Acceptable bounds but shows a minor, statistically detectable variance.
	ALP (U/L)	130.7 ±9.4	7.19%	p > 0.05	Stable; shows no signs of biliary obstruction or bone metastasis activity.
	Total Protein (g/dL)	8.4 ±0.3	3.57%	p > 0.05	Wellpreserved synthetic function.
	Albumin (g/dL)	4.2 ±0.1	2.38%	p > 0.05	Highly stable plasma protein levels.
	Globulin (g/dL)	4.2 ±0.2	4.76%	p > 0.05	Normal range; low overall cohort variability.
	A/G Ratio	1.04 ±0.04	3.85%	p > 0.05	Balanced serum protein distribution.
Renal function tests (RFT)	Urea (mg/dL)	21.6 ±1.7	7.87%	p > 0.05	Completely preserved glomerular function; low variability.
	Creatinine (mg/dL)	1.27 ±0.15	11.81%	p > 0.05	Stable clear renal clearance across the patient population.
	Sodium (Na, mmol/L)	137.0 ±1.0	0.73%	p > 0.05	Tight homeostatic control; lowest overall variation in study.
	Potassium (K, mmol/L)	4.75 ±0.20	4.21%	p > 0.05	Normal electrolyte homeostasis maintained.
Glycemic Control	HbA1c (%)	6.05 ±0.49	8.10%	p > 0.05	Wellcontrolled metabolic blood sugar level; no connection to CRP (p=0.30).

Discussion

This case series aims to highlight the longitudinal patterns of the patients' laboratory profiles, especially the consistency of their hematological, hepatic, and renal parameters. The data gathered from the patient cohort reveal the complex interplay of biological processes, especially the variables' considerable variation. The study's major goal was to measure the patients' PCA3, a prostate cancer-specific RNA, which has been proven to be effective in the detection and monitoring of prostate cancer. The PCA3 expression level ranged from 0.1 to 0.6, averaging about 0.36 and this shows that the majority of the patients had moderate expression. Expression level of PCA3 was slightly positively skewed, suggesting that the majority of patients had moderate expression of the gene, whereas others had higher expression [18].

Alshalalfa and his colleagues (2017) have noted the existence of a bimodal distribution in PCA3 levels in biopsy tissues. Low levels of PCA3 have been significantly correlated with high-grade disease, adverse pathology, metastatic progression, and poor outcomes ($p < 0.001$). However, the sensitivity of PCA3 for high-grade disease (GS≥8) was found to be low at 55%, with 42% of these having low PCA3 [19]. The difference in the expression level of PCA3 is quite consistent with what we know about prostate cancer. Prostate cancer, like all cancer types, is very individualized, meaning the cancer can be quite different from person to person with regards to the severity of the cancer and the way the tumor behaves [20-22]. The Gleason score, a measure of the tumor's differentiation and severity, ranged from 6 to 10 and most patients had a score of about 8, indicating that they had high-grade tumors and our results was similar with the results by Townsend and his colleagues (2013) by meanes of grading system [23]. Wei and his research group (2015) revealed that the high-grade tumors have been correlated to poor patient prognosis [24] and the high Gleason score and PCA3 levels indicate that the majority of the patients in this group had aggressive prostate cancer. This indicates the importance of the integration of biomarkers like PCA3 and traditional grading systems to obtain a more comprehensive patient risk assessment [25-27]. The parameters like hemoglobin, RBC, hematocrit, WBC, and platelet counts for the patients remained within the normal limits and showed no statistically significant variance ($p > 0.05$) when compared against the baseline of the control sample ($n = 21$). The hemoglobin level was between 13.5 and 14.2 g/dL, and the RBC count was about 4.7 million/mL. This indicates that the erythropoietic process was still active, and there was no major problem of anemia. The WBC and platelet counts, i.e., 6000-8000/cumm and 220,000-280,000/mm3, respectively, indicated that there was no major infection, inflammation, or blood disorder. This indicates that, despite the fact that the patients had high-grade tumors, their health status was still maintained, and this is a positive indication for the treatment and prognosis of the patients.The major observation was the fluctuation in the level of CRP, indicating systemic inflammation. Kim and his colleagues (2013) have observed that prostate cancer patient had higher CRP than benign patient [28]. The CRP values ranged from 12 to 45 mg/L, with an average value of about 24.3 mg/L and a high coefficient of variation, about 44%. Notably, these levels represent a significant elevation ($p < 0.05$) relative to the stable baseline provided by the control sample ($n = 21$). The Statistical analysis showed a positive correlation between CRP levels and the Gleason score, indicating that elevated CRP corresponds to a higher tumor grade. This suggests that physiological stress response could be an active participant in the tumor progression, aligning with previous studies that suggest these immune-mediated pathways assist prostate cancer in the advancing, metastasizing and undergoing oncogenic transformation [29-31]. Furthermore, higher CRP values correlated with the increased PCA3 expression, reinforcing idea that this systemic reactivity drives tumor advancement and may accelerate its course. These findings suggest that CRP could serve a dual purpose: as a predictive biomarker for disease progression, and as a therapeutic target where modulating this systemic response could effectively slow tumor growth [32].

All the liver function parameters of the cohort, such as bilirubin (total, direct, and indirect) as well as AST, ALT, and ALP, were generally within acceptable physiological bounds. Compared to the control sample ($n = 21$), a minor but statistically detectable variance was noted in ALT ($p <0.05$); however, the overall profile did not indicate any liver dysfunction or metastasis as observed in prostate cancer patients. The constant values of the liver enzymes indicate that the systemic effect of the prostate cancer did not affect the liver functions, which is crucial when administering chemotherapy or hormone therapy as part of the treatment regimens [33, 34]. The kidney function tests such as urea and creatinine also had constant values, averaging 21.6 mg/dL and 1.27 mg/dL, respectively, mirroring the stable renal clearance seen in the control sample ($n = 21$) with no significant difference ($p > 0.05$). The balance of the body's electrolytes, such as sodium and potassium, also had constant values within the normal range, showing good kidney function [35], with sodium demonstrating the tightest homeostatic control across the study population. Since there is no sign of kidney dysfunction, it shows that the progress of prostate cancer did not affect the kidneys as seen in patients with paraneoplastic syndromes, which is due to the effects of cancer on the body or the metastasis of cancer cells. The glycemic control levels, as reflected by HbA1c levels, were between 5.5 and 6.4 percent, showing well-controlled blood sugar levels among diabetic patients, which did not deviate significantly from the control sample ($n = 21$) values ($p > 0.05$). This indicates that the metabolic dysfunction, such as uncontrolled diabetes, did not influence progression of the disease. Currently, relationship between metabolic syndrome, prostate cancer, and systemic immune activation is being actively investigated with evidence suggesting that metabolic disruptions can significantly modulate malignancy [36-38]. Such conditions need to be monitored to understand how cancer is related to metabolic

problems and how it can be treated holistically.

Conclusion

The case series demonstrates the overall physiological stability of hematological, hepatic, and renal functions in prostate cancer patients over time, which remained remarkably consistent when compared against the healthy baseline of the control sample (n = 21), despite notable changes in molecular and inflammatory markers. Interestingly, PCA3 levels were found to vary between 0.1 and 0.6. Higher levels of PCA3 were found to be related to higher Gleason scores; hence, indicating a more aggressive disease state. Hematological parameters were found to be within normal physiological limits and did not deviate significantly from the control sample (n = 21), which shows the systemic health of the patients was well maintained. However, the results for the CRP protein showed considerable variation and a statistically significant elevation compared to the control sample (n = 21), while indicating a positive correlation with both Gleason scores and PCA3 protein levels. Furthermore, liver and kidney function tests were found to be normal and stable relative to the control group baseline, showing that no significant systemic organ dysfunction existed. This preserved organ function would help in the easier management and therapeutic tolerability of the patients. By using the advanced cluster analysis method, different types of prostate cancer patients were identified. This underscores the potential benefits that can be derived from an integrated approach to patient management, though regular longitudinal observation is necessary to obtain the maximum clinical benefits.

Consent for publication

All participants gave their informed consent before participation from the Department of Urology, AIIMS Bhopal after fulfilling the inclusion and exclusion criteria and after obtaining consent.

Declaration of conflict of interests

No.

CRediT authorship contribution statement

DA, AK, ASG and DK collected the data. SKR and SM conceptualized the manuscript. SKR wrote the first draft. SR, ShM and SM edited it. All the authors agreed and approved the final manuscript.

Author's Disclosures

The authors declare no conflicts of interest related to this work. Financial support was received from AIIMS Bhopal, India, Intramural fund (Reference Number: IHEC/SR/2024/111 Date: 01/02/2024). All authors have approved the manuscript and agree to be accountable for all aspects of the work.

Funding

AIIMS Bhopal, India, Intramural fund (AIIMSBPL/IHECSR/Nov-2024/29 dated: 23rd Dec 2024) of AIIMS Bhopal, India.

Acknowledgments

None.

References

1. Schafer EJ, Laversanne M, Sung H, Soerjomataram I, Briganti A, Dahut W, Bray F, Jemal A. Recent patterns and trends in global prostate cancer incidence and mortality: an update. Eur Urol. 2025;87(3):302–313.
2. Vlajnic T, Bubendorf L. Molecular pathology of prostate cancer: a practical approach. Pathology. 2021;53(1):36–43.
3. Huang Y, Mao J, Li X. Emerging biomarkers in prostate cancer diagnosis and treatment: insights into genetic, RNA and metabolic markers. Int J Oncol. 2026;68(2):15.
4. Lemos AEG, Matos ADR, Ferreira LB, Gimba ERP. The long non-coding RNA PCA3: an update of its functions and clinical applications as a biomarker in prostate cancer. Oncotarget. 2019;10(61):6589–6603.
5. Qin Z, Yao J, Xu L, Xu Z, Ge Y, Zhou L, Zhao F, Jia R. Diagnostic accuracy of PCA3 level in patients with prostate cancer: a systematic review and meta-analysis. Int Braz J Urol. 2020;46(5):691–704.
6. Agosti V, Munari E. Histopathological evaluation and grading for prostate cancer: current issues and crucial aspects. Asian J Androl. 2024;26(6):575–581.
7. Huang J, Reilly KH, Zhang HZ, Wang HB. Clinical evaluation of prostate cancer gene 3 score in diagnosis among Chinese men with prostate cancer and benign prostatic hyperplasia. BMC Urol. 2015;15:118.
8. Watts EL, Perez-Cornago A, Kothari J, Allen NE, Travis RC, Key TJ. Hematologic markers and prostate cancer risk: a prospective analysis in UK Biobank. Cancer Epidemiol Biomarkers Prev. 2020;29(8):1615–1626.
9. Zhou K, Li C, Chen T, Zhang X, Ma B. C-reactive protein levels could be a prognostic predictor of prostate cancer: a meta-analysis. Front Endocrinol (Lausanne). 2023;14:1111277.
10. Jensen GL, Naziri J, Hammonds KP, Jhavar SG, Swanson G. C-reactive protein is a poor marker of baseline inflammation in prostate cancer and response to radiotherapy or androgen ablation. Cureus. 2021;13(11):e19639.Stark T, Livas L, Kyprianou N. Inflammation in prostate cancer progression and therapeutic targeting. Transl Androl Urol. 2015;4(4):455-63. doi: 10.3978/j.issn.2223-4683.2015.04.12. PMID: 26816843; PMCID: PMC4708587
11. Karzai FH, Madan RA, Dahut WL. Metabolic syndrome in prostate cancer: impact on risk and outcomes. Future Oncol. 2016;12(16):1947–1955.

12. Yanase T, Kawanami T, Tanaka T, Tanabe M, Nomiyama T. Impact of metabolic disorders on prostate cancer growth: androgen and insulin resistance perspectives. Reprod Med Biol. 2017;16(3):252–257.
13. Loeb S, Curnyn C, Sedlander E. Perspectives of prostate cancer patients on Gleason scores and the new grade groups: initial qualitative study. Eur Urol. 2016;70(6):1083–1085.
14. Hamed NW, Elbeljihy HS, Hussin SA, Fouda RM, Oy EK, Magar RW. From prostate specific antigen to genomic signatures: advances in biomarkers for prostate cancer diagnosis and prognosis. Transl Oncol. 2026;66:102719.
15. Majewska Z, Zajkowska M, Pączek S, Nowiński AR, Sokólska W, Gryko M, Orywal K. The clinical relevance of tumor biomarkers in prostate cancer: a review. Cancers (Basel). 2025;17(23):3742.
16. Fernandes R, Costa C, Fernandes R, Barros AN. Inflammation in prostate cancer: exploring the promising role of phenolic compounds as an innovative therapeutic approach. Biomedicines. 2023;11(12):3140.
17. Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: examples from clinical practice. Rev Urol. 2008;10(3):175–181.
18. Alshalalfa M, Verhaegh GW, Gibb EA, Santiago-Jiménez M, Erho N, Jordan J, Yousefi K, Lam LLC, Kolisnik T, Chelissery J, Seiler R, Ross AE, Karnes RJ, Schaeffer EM, Lotan TT, Den RB, Freedland SJ, Davicioni E, Klein EA, Schalken JA. Low PCA3 expression is a marker of poor differentiation in localized prostate tumors: exploratory analysis from 12,076 patients. Oncotarget. 2017;8(31):50804–50813.
19. Malik A, Srinivasan S, Batra J. A new era of prostate cancer precision medicine. Front Oncol. 2019;9:1263.
20. Hamid AA, Sweeney CJ, Hovens C, Corcoran N, Azad AA. Precision medicine for prostate cancer: an international perspective. Urol Oncol. 2024;42(12):392–401.
21. Mytsyk Y, Borzhiyevs'kyy A, Kobilnyk Y, Shulyak AV, Dutka I, Borzhiyevs'kyy O, Górecki A. Personalized management of prostate cancer: from molecular and imaging markers to radiogenomics. Pol J Radiol. 2022;87:e58–e62.
22. Townsend NC, Ruth K, Al-Saleem T, Horwitz EM, Sobczak M, Uzzo RG, Viterbo R, Buyyounouski MK. Gleason scoring at a comprehensive cancer center: what's the difference? J Natl Compr Canc Netw. 2013;11(7):812–819.
23. Wei W, Leng J, Shao H, Wang W. High PCA3 scores in urine correlate with poor prognostic factors in prostate cancer patients. Int J Clin Exp Med. 2015;8(9):16606–16612.
24. Davalan W, Alkins R. Prognostic and predictive determinants in high-grade gliomas: integrating tumor-intrinsic biology with patient and system-level factors. Front Neurol. 2025;16:1664458.
25. Crocetto F, Musone M, Chianese S, Conforti P, Digitale Selvaggio G, Caputo VF, Falabella R, Del Giudice F, Giulioni C, Cafarelli A, Lucarelli G, Busetto GM, Ferro M, Barone B, Zattoni F, Terracciano D. Blood and urine-based biomarkers in prostate cancer: current advances, clinical applications, and future directions. J Liq Biopsy. 2025;9:100305.
26. Guo Y, Chen W, Zhu J, Wang X, Ding Y, Duan L, Mai Y, Bai Y, Niu Y, Zhao Y. Prostate cancer associated 3 (PCA3) is a potential indicator of pathological grades of prostate cancer. Discov Oncol. 2026;[Epub ahead of print].
27. Kim Y, Jeon Y, Lee H, Lee D, Shim B. The prostate cancer patient had higher C-reactive protein than BPH patient. Korean J Urol. 2013;54(2):85–88.
28. Zhou M, Liang J, Hui J, Xu J. Inflammation-related indicators have potential to improve overall quality of prostate cancer management: a narrative review. Transl Androl Urol. 2023;12(5):809–822.
29. Chen X, Li Y, Li G, Zhang X, Xie G, Huang Y, Yin H. Clinical significance of serum high sensitive C-reactive protein/albumin ratio in primary prostate biopsy. Front Oncol. 2024;14:1325524.
30. Sciarra A, Gentilucci A, Salciccia S, Pierella F, Del Bianco F, Gentile V, Silvestri I, Cattarino S. Prognostic value of inflammation in prostate cancer progression and response to therapy: a critical review. J Inflamm (Lond). 2016;13:35.
31. Du J, Lan J, Xiong J, Yang H, Xu X, Tang C, Huang G, Ying Q, Mu J, Hu Q. Efficiency of C-reactive protein in prognosis evaluation of prostate cancer: a systematic review and meta-analysis. Transl Cancer Res. 2021;10(10):4432–4439.
32. Archer M, Dogra N, Kyprianou N. Inflammation as a driver of prostate cancer metastasis and therapeutic resistance. Cancers (Basel). 2020;12(10):2984.
33. Ali S, Zehra A, Khalid MU, Hassan M, Shah SIA. Role of C-reactive protein in disease progression, diagnosis and management. Discoveries (Craiova). 2023;11(4):e179.
34. Ghoshal A, Garmo H, Hammar N, Jungner I, Malmström H, Walldius G, Van Hemelrijck M. Can pre-diagnostic serum levels of sodium and potassium predict prostate cancer survival? BMC Cancer. 2018;18(1):1169.
35. Rhee H, Vela I, Chung E. Metabolic syndrome and prostate cancer: a review of complex interplay amongst endocrine factors in pathophysiology and progression. Horm Cancer. 2016;7(2):75–83.
36. Monroy-Iglesias MJ, Russell B, Crawley D, Allen NE, Travis RC, Perez-Cornago A, Van Hemelrijck M, Beckmann K. Metabolic syndrome biomarkers and prostate cancer risk in the UK Biobank. Int J Cancer. 2021;148(4):825–834.
37. Guerrios-Rivera L, Howard LE, Wiggins EK, Hoyo C, Grant DJ, Erickson TR, Ithisuphalap J, Freedland AR, Vidal AC, Fowke JH, Freedland SJ. Metabolic syndrome is associated with aggressive prostate cancer regardless of race. Cancer Causes Control. 2023;34(3):213–221.

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

Association of CTLA-4 Gene Polymorphisms (CT60 and +49 A/G) with Hypothyroidism: A Case-Control Study in a Tertiary Care Hospital

KR Gopala Krishnan[1,2*], Hariom Sharma[1], Piyush Tailor[1], Shailesh Patel[3], D Shreshtta[4], Hina Baraiya[1]

[1]*Department of Biochemistry, Government Medical College, Bhavnagar, Gujarat*
[2]*Department of Biochemistry, ESIC Medical College and Hospital, Basaidarapur, Delhi*
[3]*Department of Biochemistry, Government Medical College, Surat, Gujarat*
[4]*Rural Health Training Centre, Najafgarh, Delhi*

Article Info

Corresponding Author:
KR Gopala Krishnan, MBBS, MD (Biochemistry)
Senior Resident
Department of Biochemistry
ESIC Medical College and Hospital, Basaidarapur, Delhi, 110015
Phone: 7338799427
E-mail: gope2011@gmail.com

Keywords

CTLA-4, Hypothyroidism, Autoimmune Thyroiditis, Genetic Polymorphism, Anti-TPO

Abstract

Background: Hypothyroidism is a major endocrine burden in India. While iodine deficiency is common, autoimmune thyroiditis is an increasingly recognized etiology. CTLA-4 gene polymorphisms are known susceptibility loci for autoimmune disorders, but data in the Indian population remain limited.

Objective: To evaluate the association of CTLA-4 exon 1 (+49 A/G) and 3' UTR (CT60) polymorphisms with hypothyroidism and autoimmune status in a Gujarati cohort.

Methods: This case-control study included 156 hypothyroid patients and 156 controls. Genotyping was performed using PCR-RFLP. Biochemical markers, including TSH and anti-TPO antibodies, were measured. Statistical analysis utilized Chi-square tests and Odds Ratios (OR) to assess genetic associations.

Results: No significant difference in age or gender was observed between groups. The CTLA-4 CT60 polymorphism showed a significant association with hypothyroidism ($\chi 2=12.491$, $p=0.002$), with the GG genotype conferring a 3.07-fold increased risk (95% CI: 1.62–5.84). Stratification by anti-TPO status revealed that the AA genotype for both CT60 and +49 A/G polymorphisms exerted a protective effect against autoimmunity (OR=0.39, p=0.032 and OR=0.22, p=0.001, respectively). Patients with the CT60 GG genotype who were anti-TPO positive exhibited significantly higher TSH levels (p=0.009).

Conclusion: CTLA-4 gene polymorphisms are significant genetic determinants of autoimmune hypothyroidism in the Gujarati population. The G allele and GG genotype are associated with increased susceptibility, while the AA genotype confers relative protection. Genotyping CTLA-4 may serve as a valuable diagnostic adjunct for risk stratification and personalized management of autoimmune thyroid disease.

Introduction

Hypothyroidism is one of the most prevalent endocrine disorders globally, characterized by the thyroid gland's inability to produce sufficient hormones to meet metabolic demands [1]. In India, the clinical burden is significant, with recent estimates suggesting an 11% prevalence rate, far exceeding figures reported in the United Kingdom or the United States [2]. While iodine deficiency historically played a primary role, Hashimoto's thyroiditis—an organ-specific autoimmune disorder—is now the leading cause of thyroid failure in iodine-sufficient regions [3].

Despite its high prevalence and the availability of standard diagnostic markers such as serum thyrotropin (TSH) and free thyroxine (T4), the pathogenesis of autoimmune hypothyroidism remains complex. It involves a sophisticated interplay of environmental triggers and genetic susceptibility [4]. Among the various candidate genes implicated in the development of autoimmune thyroid disorders, the Cytotoxic T-lymphocyte antigen-4 (CTLA-4) gene has emerged as a major susceptibility locus [5]. Located on chromosome 2q33, CTLA-4 encodes a critical negative regulator of T-cell activation, playing an essential role in maintaining peripheral self-tolerance [6].

Various polymorphisms within CTLA-4, most notably the promoter -318 C/T, exon 1 +49 A/G, and the 3' untranslated region (3' UTR) CT60, have been identified as risk factors for autoimmune endocrinopathies [7]. While several studies have linked these variants to Hashimoto's thyroiditis and Graves' disease in diverse ethnic populations, data specific to the Indian population remain limited [8]. Understanding the genetic architecture of hypothyroidism in this demographic is essential, as it may improve early risk stratification and facilitate a move toward precision medicine in clinical practice [9].

Given the current lack of localized genetic data, this study aims to evaluate the distribution of CTLA-4 (CT60 and +49 A/G) gene polymorphisms in patients with hypothyroidism attending a tertiary care hospital in Bhavnagar. By correlating these genotypes with clinical variables, this research seeks to establish a foundational genetic database to aid in the identification of individuals with high susceptibility to thyroid dysfunction.

Methodology

Patient Recruitment

The study was conducted in the Department of Biochemistry, Government Medical College, Bhavnagar. Informed consent was taken from all the patients who participated in this study. A total of 156 Group I: Hypothyroidism (Case) and 156 Group II: Non-Hypothyroidism (Control) were included in the present study. The study protocol was approved by the Institutional Ethics Committee of the Government Medical College, Bhavnagar (Approval No: 1201/2022). This was a case control type and was conducted from the period of October 2022 to September 2023.

Inclusion and exclusion criteria

Patients were eligible for inclusion if they were between 25 and 75 years of age and had a confirmed clinical diagnosis of primary hypothyroidism. To ensure the study focused specifically on primary hypothyroidism, several exclusion criteria were applied. Subjects were excluded if they presented with secondary or tertiary hypothyroidism resulting from pituitary or hypothalamic dysfunction. Furthermore, patients with a history of thyroid surgery, pregnancy, or those requiring the continuous or periodic use of corticosteroids were excluded from the study. We also excluded individuals with severe comorbidities, such as active malignancies or severe systemic infections, which could influence TSH levels or immunological status. Rather than excluding autoimmune disease, our cohort intentionally included patients with Hashimoto's thyroiditis—identified by elevated anti-TPO antibody titers - to allow for a comprehensive evaluation of the association between CTLA-4 gene polymorphisms and autoimmune-mediated thyroid dysfunction.

Diagnostic Criteria for Hypothyroidism

The diagnosis of primary hypothyroidism is typically based on the demonstration of an elevated Serum Thyroid-Stimulating Hormone (TSH) level in conjunction with a low Serum Free Thyroxine (FT4) level. Subclinical Hypothyroidism: Characterized by an elevated TSH with a normal FT4 level. Overt Hypothyroidism: Characterized by an elevated TSH with a low FT4 level. The most widely accepted guidelines used in clinical research are those published by the American Thyroid Association (ATA) and the European Thyroid Association (ETA) [10,11].

Participants' Examination and Measurements

Patients were interviewed for medical history. The present and past history of each case were recorded in detail regarding their general information i.e. name, age, sex, address, occupation, economic status, personal habits especially smoking, medication, and history suggestive of any systemic illness especially hypothyroidism. The history of Hypothyroidism was based on the patient's self-report of a prior physician diagnosis and was under treatment with any drug.

Sampling and Biochemical Analysis

After completing the initial history taking, 5ml of peripheral venous blood was collected by vein puncture from each of the subjects. In Plain & EDTA vacutainer for TSH, Free T3 , Free T4, Anti TPO test were done by chemiluminescence based assay kit on fully automated analyser and EDTA used for isolation of DNA.

DNA Extraction

Genomic DNA was extracted from peripheral blood leukocytes collected in EDTA-anticoagulated tubes using the HiPurA™

Blood Genomic DNA Miniprep Purification Kit (HiMedia Laboratories, India), based on the selective binding of DNA to a silica-based membrane in the presence of chaotropic salts. The procedure was performed in accordance with the manufacturer's protocol. Briefly, 200 µL of whole blood was transferred to a collection tube and treated with 20 µL of Proteinase K (20 mg/mL) and 20 µL of RNase A solution (20 mg/mL), followed by brief vortexing. The sample was incubated at room temperature for 2 minutes to ensure enzymatic digestion. Subsequently, 200 µL of lysis buffer was added, the mixture was vortexed to achieve homogeneity, and incubated at 55°C for 10 minutes. Following the lysis reaction, 200 µL of absolute ethanol (96–100%) was added to the lysate to facilitate binding conditions.

The resulting mixture was loaded onto a HiElute™ Miniprep Spin Column and centrifuged at 10,000 rpm for 1 minute; the flow-through was discarded. The column underwent a two-stage washing process: first, with 500 µL of pre-wash solution (centrifuged at 10,000 rpm for 1 minute), and second, with 500 µL of wash solution (centrifuged at 13,000 rpm for 3 minutes). To ensure the complete removal of residual ethanol, an additional centrifugation step was performed for 1 minute at the same speed. Finally, the genomic DNA was eluted in two sequential 100 µL aliquots of Elution Buffer (10 mM Tris-Cl, pH 8.5) by incubating for 1 minute at room temperature followed by centrifugation at 6,500 x g. The extracted DNA was assessed for quality and subsequently stored at -20°C until further amplification, which was completed within 72 hours of extraction.

Concentration, yield and purity of DNA

Concentration and the purity of the genomic DNA in each sample were accessed in IMPLEN nanophotometer p300. Elution Buffer (ET) to dilute samples and to calibrate the nanophotometer, absorbance was measured at 260 nm, 280 nm, and 320 nm. The 320 nm absorbance was used to correct for background absorbance. An absorbance of 1.0 at 260 nm corresponds to approximately 50 µg/ml of DNA. The A260 - A320 /A280 -A320 ratio should be 1.6-1.9. Purity was determined by calculating the ratio of absorbance at 260 nm to absorbance at 280 nm.

Genetic Polymorphism of CTLA4 gene

The CTLA-4 gene polymorphisms (+49 A/G and 3' UTR CT60) were genotyped using the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique. Amplification was performed in a Bio-Rad T100™ Thermal Cycler using the HiMedia 2X PCR Taq Master Mix. The final 25 µL reaction volume consisted of 12.5 µL of 2X Taq Master Mix, 0.4 µL of each forward and reverse primer (100 ng/µL), 1 µL of genomic DNA template (50 ng/µL), and 10.7 µL of nuclease-free water. The primer sequences and conditions were selected based on validated protocols for human genomic DNA analysis. SNP/Gene CTLA4 +49A/G primers, forward: 5' AAGGCTCAGCTGAACCTGGT 3' and reverse: 5' CTTTGCCTTATTTGCTGCCGC 3' sequences and SNP/Gene CTLA4 CT60 A/G primers, forward: 5' CACCACTATTTGGGATATACC 3' and reverse: 5' AGCTCTATATTTCAGGAAGGC 3' sequences were employed. The thermal cycling profile included an initial denaturation at 95°C for 10 minutes, followed by 30 cycles of denaturation (95°C for 30 seconds), annealing (60°C for 30 seconds), and extension (72°C for 30 seconds), with a final extension at 72°C for 5 minutes. Following amplification, 15 µL of the PCR products were digested for 16 hours at 37°C using 5 U of the specific restriction endonucleases BstEII for the +49 A/G polymorphism and NcoI for the CT60 polymorphism. For the +49 A/G polymorphism, digestion of the 271 bp PCR product with BstEII yielded distinct banding patterns: the A allele was characterized by cleavage into 249 bp and 22 bp fragments, whereas the G allele remained undigested at 271 bp. Similarly, for the 3' UTR CT60 A/G polymorphism, digestion of the 216 bp PCR product with NcoI allowed for genotype discrimination: the A allele was identified by fragments of 174 bp and 42 bp, while the G allele remained as an undigested 216 bp band. Based on these banding patterns, three genotypes were identified for both polymorphisms via 2% agarose gel electrophoresis: homozygous wild-type (AA), heterozygous (AG), and homozygous variant (GG). All samples were scored independently to ensure accuracy, and a subset of samples was re-analyzed to confirm the reproducibility of the genotyping results (Figure 1 and 2).

Figure 1: 49AG 271,249,22 bp fragment (lane no. 3,6) suggestive of AG genotype and 271 bp fragment (lane No.1,2,4) suggestive of GG genotype and 249,22 bp fragment (lane No.5,7) suggestive of AA genotype, distinguished using ethidium bromide under UV (ultra-violet) light in Gel Document system.

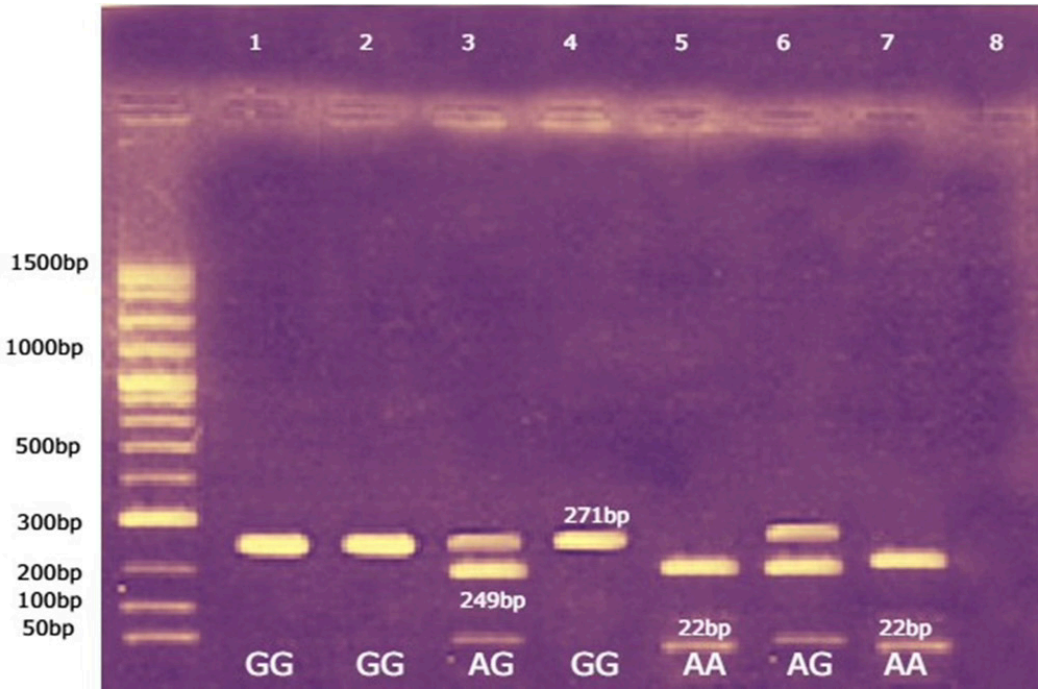
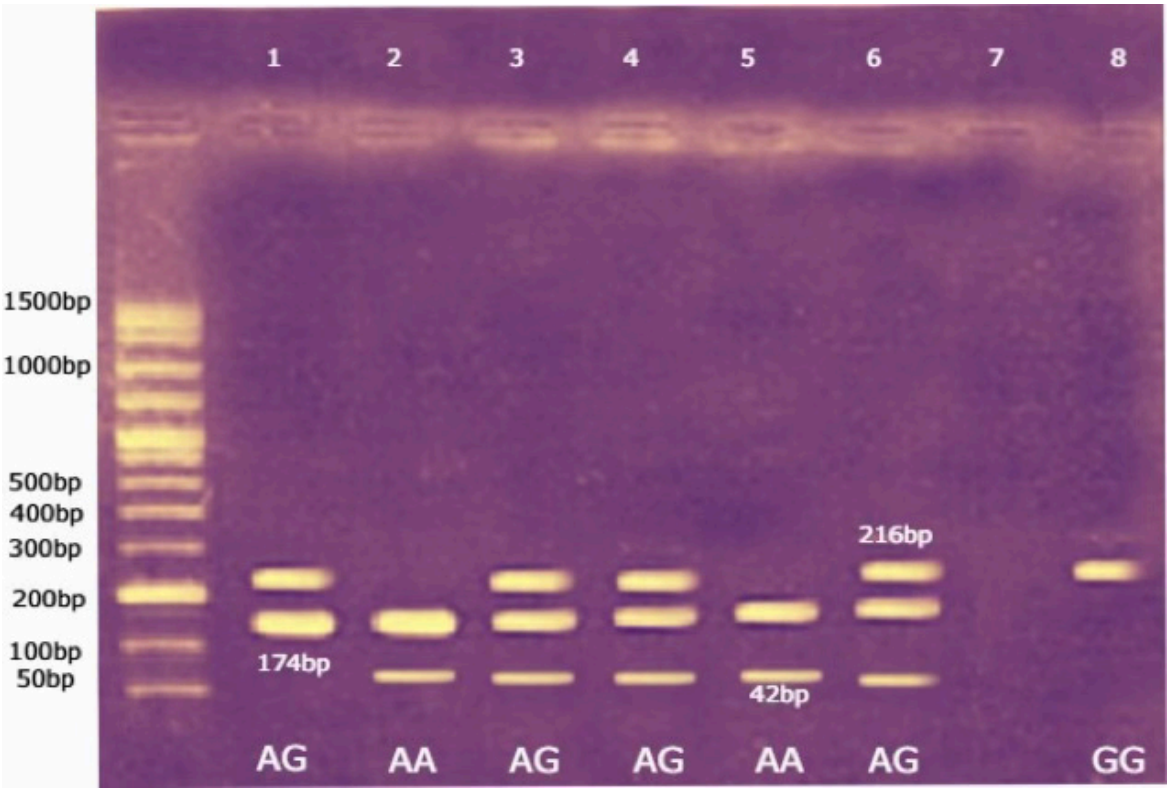


Figure 2: CT60 216,174,42 bp fragment (lane No.1,3,4,6) suggestive of AG genotype 216, bp fragment (lane no. 8) suggestive of GG genotype and 174,42 bp fragment (lane No.2,5) suggestive of AA genotype, distinguished using ethidium bromide under UV (ultra-violet) light in Gel Document system.



Data analysis

All data were analyzed using web-based statistical platforms (SPSS 25 version). Continuous variables were expressed as mean ± standard deviation (SD) and compared between groups using the student t-test. Categorical variables were evaluated using the Chi-square test. To determine whether the allele distribution within the study groups adhered to the Hardy-Weinberg equilibrium, the chi2 goodness-of-fit test was applied. The association between CTLA-4 gene polymorphisms and hypothyroidism was assessed by calculating the Odds Ratio (OR) with a 95% confidence interval (CI). All probability values were two-sided, with $p<0.05$ considered statistically significant, $p<0.001$ as highly significant, and $p>0.05$ as not significant.

Result

Demographic and biochemical status

The baseline characteristics of the study population are summarized in Table 1. There was no statistically significant difference in mean age between the hypothyroid group (39.89±12.38 years) and the control group (41.19±14.89 years; p=0.402). The gender distribution was comparable across both groups, with a higher prevalence of females (70.5% in Group I vs. 72.4% in Group II). As expected, biochemical analysis revealed a highly significant difference ($p<0.001$) in thyroid profile markers, with elevated TSH and anti-TPO levels, and lower FT3 and FT4 levels observed in the hypothyroid cohort compared to the control group.

Table 1: Demographic and Clinical Characteristics of the Study Population.

Parameter	Group I: Hypothyroidism (n=156)	Group II: Controls (n=156)	pvalue
Age (years), Mean ± SD	39.89 ± 12.38	41.19 ± 14.89	0.402
Gender, n (%)			
Male	46 (29.49)	43 (27.56)	0.692
Female	110 (70.51)	113 (72.44)	
Biochemical Markers			
TSH (µIU/mL)	19.03 ± 25.78	1.97 ± 0.90	<0.001*
FT3 (pg/mL)	1.90 ± 0.84	2.88 ± 0.30	<0.001*
FT4 (ng/dL)	0.95 ± 0.73	1.62 ± 0.38	<0.001*
AntiTPO (IU/mL)	31.12 ± 16.44	17.70 ± 5.11	<0.001*

*The Chi-square test was used to calculated the categorial data (gender) and the student t-test was used to calculate the p-value for continuous data. *p<0.05 was considered as statistically significant.*

CTLA-4 Genotypes on Thyroid Hormone Profiles in Hypothyroid Patients (Group I)

We analyzed the correlation between CTLA-4 genotypes and thyroid hormone profiles stratified by anti-TPO status. Notably, in patients carrying the CT60 GG genotype, there was a statistically significant difference in mean TSH levels between anti-TPO positive and negative cohorts (p=0.009), with anti-TPO positive individuals exhibiting markedly higher TSH concentrations. No significant associations were observed between the +49 A/G genotypes and thyroid hormone levels, nor were there significant correlations identified for FT3 or FT4 levels across any of the studied genotypes (Table 2).

Table 2: Influence of CTLA-4 Genotypes on Thyroid Hormone Profiles in Hypothyroid Patients (Group I).

Polymorphism	Genotype	TSH (µIU/mL) (n=105) (TPOve)	TSH (µIU/mL) (n=51) (TPO+ve)	pvalue	FT3 (pg/mL) (n=105) (TPOve)	FT3 (pg/mL) (n=51) (TPO+ve)	pvalue
+49 A/G	AG	16.33	25.64	0.350	1.54	1.74	0.659
	GG	23.47	24.85	0.199	1.87	2.12	0.480
	AA	15.69	8.45	0.086	2.02	1.89	0.480
CT60	AG	20.35	27.49	0.568	1.75	1.90	0.741
	GG	15.98	27.59	**0.009***	1.93	1.90	0.659
	AA	16.34	7.15	0.062	2.01	1.72	0.659

*The student t-test was used to calculate the p-value for continuous data. *p<0.05 was considered as statistically significant.*

Genotype and Allele Distribution of CTLA-4 Polymorphisms

The genotype distribution and allele frequencies of the CTLA-4 +49 A/G and CT60 polymorphisms are summarized in Table 3. The genotype frequencies for the +49 A/G polymorphism did not show a statistically significant difference between the hypothyroid group and the control group ($\chi2=3.510$, p=0.173). The A allele frequency was 57% in the hypothyroid group compared to 68% in the control group, while the G allele frequency was 42% and 31%, respectively.

In contrast, the CT60 polymorphism revealed a significant association with hypothyroidism ($\chi2=12.491$, p=0.002). While the distribution of the AA genotype was lower in the hypothyroid group (63 vs. 83), the homozygous variant GG genotype was markedly more prevalent in the hypothyroid group (42 vs. 18). Consequently, individuals carrying the CT60 GG genotype exhibited a significantly increased risk of hypothyroidism (OR = 3.07; 95% CI: 1.62–5.84). The allele frequencies for CT60 further supported this finding, with the G allele frequency being 44% in the hypothyroid group versus 29% in the control group.

Table 3: Genotype and Allele Frequencies of CTLA-4 Polymorphisms in Hypothyroid.

Polymorphism	Genotype	Group I (Hypothyroidism)		Group II (Non-Hypothyroidism)		χ2 d.f.2	P value d.f.2	OR	OR(95% CI)	
									Lower Bound	Upper Bound
49A/G	AA	67		78		3.510	0.1729	0.87	0.53	1.43
	AG	56		57				1.14	0.7	1.87
	GG	33		21				1.83	0.97	3.46
	Allele	A	G	A	G					
	Frequency	0.57%	0.42%	0.68%	0.31%					
CT60	AA	63		83 55 18		12.491	0.0019*	0.82	0.5	1.35
	AG	51						1.22	0.74	2.02
	GG	42						3.07	1.62	5.84
	Allele	A	G	A	G					
	Frequency	0.55%	0.44%	0.70%	0.29%					

Chi-square test; df: degrees of freedom; OR: Odds Ratio; CI: Confidence Interval. p<0.05 indicates statistical significance. The risk (OR) for genotypes AG and GG was calculated relative to the homozygous wild-type (AA) genotype.

Association of CTLA-4 Polymorphisms with Autoimmune Status

The association between CTLA-4 genotype distributions and autoimmune status (anti-TPO positive vs. negative) in patients with hypothyroidism is presented in Table 4. Using the GG genotype as the reference, we evaluated the risk profile for variant genotypes. For the CT60 polymorphism, the AA genotype was significantly underrepresented in the autoimmune (TPO positive) group compared to the TPO-negative group (OR = 0.40; 95% CI: 0.16–0.98, p<0.05). For the +49 A/G polymorphism, the AA genotype demonstrated a highly significant protective effect against autoimmune thyroiditis compared to the GG reference group (OR = 0.22; 95% CI: 0.09–0.57, p<0.001). These findings indicate that while the GG genotype serves as a risk marker for autoimmunity in this population, the AA genotype potentially confers a degree of genetic resistance to the development of autoimmune thyroid antibodies (Table 4).

Table 4: Association of CTLA-4 Polymorphisms with Autoimmune Status (Anti-TPO Positive) in Hypothyroid Patients.

Polymorphism	Genotype	(TPO+; n=51)	(TPO; n=105)	OR (95% CI)	pvalue
CT60	GG (Ref)	18 (35.29%)	26 (24.76%)	1.00	
	AG	20 (39.22%)	31 (29.52%)	0.93 (0.41–2.12)	0.866
	AA	13 (25.49%)	48 (45.71%)	0.39 (0.17–0.92)	**0.032***
+49 A/G	GG (Ref)	17 (33.33%)	21 (20.00%)	1.00	
	AG	25 (49.02%)	31 (29.52%)	0.99 (0.43–2.28)	0.992
	AA	9 (17.65%)	53 (50.48%)	0.22 (0.08–0.54)	**0.001***

*Data are presented as n (%). TPO+: Anti-TPO positive (Autoimmune); TPO-: Anti-TPO negative. OR: Odds Ratio calculated using the GG genotype as the reference group. CI: 95% Confidence Interval. *Significant at p<0.05; Highly significant at p<0.001.*

Discussion

The present study showed that low iodine intake is not the sole etiological candidate for hypothyroidism in India, as approximately one-third of the patients included in this study tested positive for anti-TPO antibodies. This trend aligns with recent large-scale epidemiological studies across India, which identified anti-TPO positivity as a primary indicator for thyroid disease, though those studies did not specifically isolate the prevalence of autoimmune hypothyroidism [12]. No relationship between the age of the patient and the presence of anti-TPO antibodies was found, a finding consistent with previous longitudinal research [13].

Our study provides evidence that CTLA-4 gene polymorphisms—specifically exon 1 (+49 A/G) and the 3' UTR (CT60)—are significant susceptibility loci for autoimmune hypothyroidism in the Gujarati population. We observed that the G allele at both loci correlates with increased disease risk, whereas the AA genotype consistently appears to confer a degree of genetic protection.

The association between CTLA-4 variants and hypothyroidism is multifaceted. We found that the CT60 GG genotype is a robust marker of susceptibility, demonstrating a 3.07-fold increased risk of disease (p=0.0019, OR=3.07). This finding is consistent with literature from Caucasian and Japanese populations, suggesting that the CT60 locus serves as a conserved marker for autoimmune thyroid disease across diverse ethnic groups [14, 15]. Furthermore, our analysis of thyroid profiles revealed that among hypothyroid patients, those carrying the CT60 GG genotype who are anti-TPO positive exhibit significantly higher TSH concentrations (p=0.009) compared to their anti-TPO negative counterparts. This suggests that these genetic variants not only predispose individuals to the disease but may also exacerbate the severity of the hypothyroid phenotype in the presence of an active autoimmune response [16].

A critical observation in our study is that while iodine deficiency is a known etiological factor in India, it is not the sole driver of hypothyroidism. Approximately one-third of our hypothyroid patients were anti-TPO positive, highlighting the significant burden of autoimmune-mediated thyroid failure. In our cohort, we found no significant relationship between patient age, gender, and the presence of anti-TPO antibodies (p=0.299), suggesting that genetic predisposition—rather than demographic factors—is a primary driver of the autoimmune component [17].

Our data further elucidates the role of CTLA-4 haplotypes. We identified that the AG and GG haplotypes for the +49 A/G and CT60 sites were more frequently observed in anti-TPO positive patients. Specifically, the G allele at both sites was associated with disease susceptibility, corroborating findings in Japanese

and other global populations [18, 19]. However, the presence of autoimmune hypothyroidism in some patients with "protective" AA genotypes, and the absence of disease in some individuals with "risk-associated" GG genotypes, reinforces the conclusion that autoimmune hypothyroidism is a complex, multifactorial disorder [20]. These discrepancies likely reflect the interplay between genetic susceptibility, epigenetic modifications, and environmental triggers, such as dietary iodine intake or other lifestyle factors.

By closely matching our study and control groups for age and sex, we were able to minimize confounding variables, allowing for a clearer association between CTLA-4 genotypes and thyroid dysfunction. While our findings support the status of CTLA-4 on chromosome 2q33 as a critical susceptibility locus [21, 22], they also emphasize the need to look beyond single-gene models to fully understand the pathogenesis of autoimmune thyroid disease.

Strengths and Limitations

The primary strength of this study lies in the rigorous matching of the hypothyroid group and controls for age, sex, and clinical parameters, which minimizes confounding variables. Furthermore, by stratifying patients by anti-TPO status, we were able to specifically isolate the genetic contribution to the autoimmune component of hypothyroidism.

The study has some limitations. First, the sample size, while adequate for a preliminary regional study, may not be sufficient to detect the influence of rare variants or low-frequency haplotypes. Second, as a hospital-based case-control study in a tertiary care center, there is a potential for selection bias. Finally, while we have established a genetic association, functional analysis to determine how these polymorphisms alter CTLA-4 protein expression or T-cell regulatory capacity in this specific population was outside the scope of this study and warrants future investigation.

Conclusion

This study concludes that CTLA-4 gene polymorphisms are significant genetic determinants of autoimmune hypothyroidism in the studied population. Specifically, the G allele and the GG genotype at the +49 A/G and CT60 loci are associated with increased susceptibility to the disease and the development of anti-TPO antibodies. These findings support the role of CTLA-4 as a critical susceptibility locus for autoimmune thyroid pathogenesis. We propose that CTLA-4 genotyping could serve as a valuable diagnostic adjunct for risk stratification and the early identification of individuals predisposed to autoimmune thyroid dysfunction, potentially facilitating personalized management approaches.

Statements and Disclosures

CRediT Author Statement

KR Gopala Krishnan: Conceptualization, Methodology, Writing – Original Draft.
Hariom Sharma: Formal analysis.
Piyush Tailor: Resources, Project administration.
Shailesh Patel: Formal analysis
D Shreshtta: Software, Formal analysis.
Hina Baraiya: Writing – Review & Editing.
As the corresponding author, Dr. KR Gopala Krishnan ensures that these descriptions are accurate and have been agreed upon by all authors.

Declaration of Conflicting Interests

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Ethical Approval and Compliance

This study protocol was approved by the Institutional Ethics Committee of the Government Medical College, Bhavnagar (Approval No: 1201/2022). All procedures involving human subjects were conducted in strict compliance with the ethical principles outlined in the 1964 Declaration of Helsinki and its later amendments. Written informed consent was explicitly obtained from all individual patients prior to their enrollment in the study.

Artificial Intelligence (AI) Statement

The authors declare that an AI tool was utilized only for refine technical language formatting. The human authors reviewed, edited, and take full responsibility for the final text and intellectual content of this manuscript.

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 genetic and biochemical datasets analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. Chiovato L, Magri F, Carlé A. Hypothyroidism in context: where we've been and where we're going. Adv Ther. 2019;36(Suppl 2):47-58. Available from: https://pubmed.ncbi.nlm.nih.gov/31485975/ (accessed: 05/06/2026).
2. Unnikrishnan AG, Menon UV. Thyroid disorders in India: An epidemiological perspective. Indian J Endocrinol Metab. 2011;15(Suppl 2):S78-81. Available from: https://pubmed.ncbi.nlm.nih.gov/21966515/ (accessed: 08/06/2026).
3. Caturegli P, De Remigis A, Rose NR. Hashimoto thyroiditis: clinical and diagnostic criteria. Autoimmun Rev. 2014;13(4-5):391-397. Available from: https://pubmed.ncbi.nlm.nih.gov/24434360/ (accessed: 05/06/2026).
4. Tomer Y, Huber A. The etiology of autoimmune thyroid

disease: a story of genes and environment. J Autoimmun. 2009;32(3-4):231-239. Available from: https://pubmed.ncbi.nlm.nih.gov/19307093/ (accessed: 12/06/2026).
5. Tomer Y. The genetics of autoimmune thyroid disease. Endocrinol Metab Clin North Am. 2009;38(2):331-348. Available from: https://pubmed.ncbi.nlm.nih.gov/19446217/ (accessed: 12/06/2026).
6. van Kooten C, Boon L, Paanen-Visser J, et al. CTLA-4: A checkpoint in the immune system. Immunol Lett. 2017;186:13-20. Available from: https://pubmed.ncbi.nlm.nih.gov/28325732/ (accessed: 12/06/2026).
7. Ueda H, Howson JM, Esposito L, et al. Association of the CTLA-4 gene with autoimmune disease. Nature. 2003;423(6939):506-511. Available from: https://pubmed.ncbi.nlm.nih.gov/12717450/ (accessed: 15/06/2026).
8. Banerjee A, Saxena A, Sinha N, et al. Association of CTLA-4 gene polymorphisms with autoimmune thyroid disease in an Indian population. J Clin Endocrinol Metab. 2008;93(5):1848-1852. Available from: https://pubmed.ncbi.nlm.nih.gov/18319323/ (accessed: 15/06/2026).
9. Vanderpump MP. The epidemiology of thyroid disease. Br Med Bull. 2011;99:39-51. Available from: https://pubmed.ncbi.nlm.nih.gov/21831967/ (accessed: 15/06/2026).
10. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. Endocr Pract. 2012;18(6):988-1028. Available from: https://pubmed.ncbi.nlm.nih.gov/23246686/ (accessed: 15/06/2026).
11. Pearce SH, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. Eur Thyroid J. 2013;2(4):215-228. Available from: https://pubmed.ncbi.nlm.nih.gov/24455467/ (accessed: 15/06/2026).
12. Unnikrishnan AG, Kalra S, Sahay RK, et al. Thyroid epidemiology in India: A multi-city study. Indian J Endocrinol Metab. 2013;17(3):473-475. Available from: https://pubmed.ncbi.nlm.nih.gov/23869304/ (accessed: 15/06/2026).
13. Vanderpump MP, Tunbridge WM, French JM, et al. The Whickham survey: 20-year follow-up. Clin Endocrinol (Oxf). 1995;43(1):55-68. Available from: https://pubmed.ncbi.nlm.nih.gov/7641412/ (accessed: 15/06/2026).
14. Tomer Y. Genetics of autoimmune thyroid disease: a comprehensive review. J Autoimmun. 2014;53:1-12. Available from: https://pubmed.ncbi.nlm.nih.gov/24856545/ (accessed: 15/06/2026).
15. Kotsa K, Polyzos SA, Gazouli M, et al. CTLA-4 gene polymorphisms in autoimmune thyroid disease. Thyroid. 2007;17(5):455-461. Available from: https://pubmed.ncbi.nlm.nih.gov/17542699/ (accessed: 15/06/2026).
16. Chistiakov DA. CTLA-4 and autoimmune diseases. Immunol Lett. 2008;117(1):1-10. Available from: https://pubmed.ncbi.nlm.nih.gov/18243354/ (accessed: 15/06/2026).
17. Zaletel K, Gaberšček S, Pirnat E, et al. The influence of age and gender on the prevalence of thyroid autoimmunity. Thyroid. 2011;21(3):289-294. Available from: https://pubmed.ncbi.nlm.nih.gov/21190448/ (accessed: 15/06/2026).
18. Akbarzadeh S, Miresmaeili SM, Gholamrezaei MS, et al. CTLA-4 gene polymorphisms in Iranian patients with autoimmune thyroid disease. Endocr J. 2012;59(1):1-7. Available from: https://pubmed.ncbi.nlm.nih.gov/22033560/ (accessed: 15/06/2026).
19. Jin Y, Wang L, Zhang S, et al. Association of CTLA-4 gene polymorphisms with AITD: a meta-analysis. Endocrine. 2013;43(3):570-580. Available from: https://pubmed.ncbi.nlm.nih.gov/22996160/ (accessed: 15/06/2026).
20. Tomer Y, Huber A. The etiology of autoimmune thyroid disease. J Autoimmun. 2009;32(3-4):231-239. Available from: https://pubmed.ncbi.nlm.nih.gov/19307093/ (accessed: 15/06/2026).
21. Banerjee A, Saxena A, Sinha N, et al. CTLA-4 and thyroid autoimmunity. J Clin Endocrinol Metab. 2008;93(5):1848-1852. Available from: https://pubmed.ncbi.nlm.nih.gov/18319323/ (accessed: 15/06/2026).
22. Noh JY, Sato A, Suzuki K, et al. CTLA-4 as a susceptibility gene for autoimmune thyroid disease. Endocr J. 2010;57(8):723-729. Available from: https://pubmed.ncbi.nlm.nih.gov/20595785/ (accessed: 15/06/2026).

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

Evaluation of Infectious Disease Immunoassays for HIV 1/2 Ab-Ag Combo and Syphilis Ab on Alinity i, Maglumi X3 and CL-1200i CLIA Systems

Diana G. Coste[1*], Stefan Dascalu[1,2], Anthony J. Nagle[1]

[1]*Dig Worldwide Ltd, Innovation House, Sandwich, United Kingdom*
[2]*Department of Biology, University of Oxford, Oxford, United Kingdom*

Article Info

[*]***Corresponding Author:***
Diana G. Coste
Dig Worldwide Ltd, Innovation House, Sandwich, United Kingdom
E-mail: dgc@dig-worldwide.com

Keywords

Analytical sensitivity, Chemiluminescent immunoassay, CLIA, Clinical sensitivity, Clinical specificity, Early infection detection, HIV 1/2 Combo Ab/Ag, HIV infection, Infectious disease serology, Performance evaluation study, Seroconversion, Sexually transmitted disease control, Syphilis Anti-TP, Transfusion screening

Abstract

Background: To support global HIV and syphilis elimination, highly sensitive serological screening is essential. Early detection is critical, as evolving diagnostic landscapes can alter antigen-antibody dynamics. Independent evaluations are necessary to assess the real-world performance of 4th generation assays in detecting low-level biomarkers during the seroconversion window.

Objectives: Performance evaluation and comparison of HIV and syphilis infectious disease serology screening across multiple analytical performance studies conducted on three commercially available automated chemiluminescent immunoassay (CLIA) systems.

Materials & Methods: Performance of the HIV 1/2 Ab-Ag Combo and Syphilis immunoassays was evaluated across three CLIA systems: Abbott Alinity i (Abbott Laboratories Inc., Abbott Park, Illinois, USA), Snibe Maglumi X3 (Shenzhen New Industries Biomedical Engineering Co., Ltd., Shenzhen, China), and Mindray CL-1200i (Mindray Bio-Medical Electronics Co., Ltd., Shenzhen, China). HIV analytical sensitivity was assessed using the 1st WHO International Reference Panel (NIBSC 16/210), representing 10 diverse subtypes and Group O. Clinical performance was determined using 132 specimens from 10 seroconversion panels, 15 HIV subtype samples, and 1,500 negative patient samples. For syphilis, clinical sensitivity and specificity were established using 250 confirmed positive and 250 negative clinical specimens. All testing followed ISO 15189:2022 and CLSI guidelines.

Results: All systems demonstrated acceptable analytical precision for routine clinical use. For the HIV reference panel, the Alinity i consistently exhibited the highest analytical sensitivity across all 10 subtypes and Group O (mean: 0.728 IU/mL; range: 0.658-0.764 IU/mL), followed by Maglumi X3 (mean: 0.819 IU/mL; range: 0.790-0.880 IU/mL) and CL-1200i (mean: 2.313 IU/mL; range: 2.05-2.60 IU/mL). These analytical differences corresponded to distinct variations in early detection windows across the

evaluated seroconversion panels; the mean time until the first reactive result was 11.8 days (median: 9.0; range: 0-35) for Alinity i (59.1% overall detection), 15.8 days (median: 13.5; range: 0-37) for Maglumi X3 (46.9% detection), and 19.1 days (median: 16.0; range: 7-44) for CL-1200i (34.8% detection). Clinical specificity for HIV was 99.87% (1498/1500) for Alinity i, while both Maglumi X3 and CL-1200i demonstrated a higher number of false reactive results (n=5 for both vs. n=2 for Alinity i), resulting in a lower specificity of 99.67% (1495/1500). For syphilis, clinical sensitivity was 99.60% for Alinity i and 99.20% for Maglumi X3, whereas CL-1200i showed a diagnostic gap at 96.00%. Syphilis specificity was 99.20% for Alinity i and Maglumi X3, compared to 97.20% for CL-1200i.

Conclusions: Per the In Vitro Diagnostic Regulation (IVDR), post-market evaluations of infectious disease immunoassays are increasingly critical. Our findings underscore this necessity by revealing significant analytical and clinical performance differences among widely used CLIA systems. These variations carry direct implications for the reliable detection of early infections, and suggest important hypotheses regarding the potential operational impact on the diagnosis and monitoring of individuals receiving PrEP and early ART. In these increasingly common clinical scenarios, even minor analytical differences between assays represent potential diagnostic gaps that warrant careful consideration within real-world testing workflows.

Introduction

Human immunodeficiency virus (HIV) and syphilis remain significant global public health concerns, posing ongoing risks to blood safety and maternal-fetal health [1–4]. HIV targets the immune system, progressively weakening host defences against opportunistic infections and certain cancers. The global burden of HIV remains staggering; by the end of 2024, approximately 40.8 million people were living with the virus, with an estimated 1.3 million new infections occurring that year [1]. The virus is primarily transmitted through the exchange of bodily fluids via sexual contact, unsafe blood transfusions, and from mother to child. Syphilis is a curable sexually transmitted infection (caused by the bacterium Treponema pallidum) that is experiencing a global resurgence [1,5]. Recent data indicates that approximately 8 million adults acquire syphilis annually [1,2,6]. Like HIV, syphilis is transmitted through sexual contact, unsafe blood practices, and during pregnancy. The impact of maternal syphilis is particularly devastating [7], resulting in over hundreds of thousands of congenital syphilis cases worldwide [8], causing significant adverse birth outcomes globally [9]. Furthermore, syphilis infection is consistently associated with an increased risk of HIV acquisition. Despite the availability of effective prevention and treatment strategies [9], both infections persist across diverse populations, including blood donors, pregnant women, and, although rare, individuals intermittently utilizing pre-exposure prophylaxis (PrEP) [2,10]. Therefore, highly accurate and reliable laboratory screening remains essential for timely diagnosis and public health prevention and control.

Serological testing represents the primary entry point for diagnosing HIV and syphilis in most screening algorithms. However, early infections are frequently asymptomatic and may feature low or delayed antibody responses. For HIV, this diagnostic challenge is increasingly relevant in the context of expanded PrEP utilization and early antiretroviral therapy (ART) initiation [11]. While these interventions are critical for patient care, exposure to antiretrovirals can significantly suppress or alter typical serological dynamics, potentially causing false-negative or borderline results during acute infection. Furthermore, the combination of these blunted antibody responses and the broad diversity of circulating HIV subtypes necessitate highly sensitive assays capable of robust performance across complex clinical profiles [12]. Ultimately, diagnostic delays caused by assay limitations can postpone confirmatory nucleic acid testing, treatment optimization, and critical partner notification.

Since syphilis cannot be reliably distinguished from other conditions on clinical grounds alone, laboratory confirmation is vital [11]. Test interpretation can be complicated by the infection stage, the host immune response, and physiological states such as pregnancy. Untreated syphilis can last for years and cause severe health issues. Missed or delayed detection is particularly dangerous for pregnant women; inadequate or delayed treatment can result in congenital infection in 50% to 80% of cases, leading to adverse outcomes such as stillbirth, neonatal death, prematurity, and low birth weight [11].

Automated chemiluminescent immunoassay (CLIA) systems are widely used for high-throughput screening. However, differences in assay design and signal detection can significantly affect their analytical and clinical performance [13]. Therefore, evaluations under routine laboratory conditions are necessary to characterize real-world performance beyond manufacturer claims.

To address these diagnostic challenges, our study presents a comparative performance evaluation of HIV 1/2 Ab-Ag Combo and Syphilis immunoassays across three widely used automated CLIA platforms: the Abbott Alinity i, Snibe Maglumi X3 and Mindray CL-1200i. Conducted under routine clinical laboratory conditions, our assessment aims to characterize the real-world analytical and clinical performance of these CLIA devices beyond manufacturer claims. Ultimately, our study seeks to reveal if any practical differences in assay capabilities are present, thus providing evidence-based insights which can guide system selection for reliable screening.

Materials and Methods

Study design

This study was conducted under routine clinical laboratory conditions to reflect daily diagnostic conditions rather than controlled manufacturer validation environments. The study

protocol, objectives and the materials used are summarized in Table 1. All measurements were performed by trained laboratory personnel using standard operating procedures, routine calibration and quality control practices, and manufacturer recommended assay settings. Residual serum samples were prospectively collected after completion of routine diagnostic testing and were fully de-identified prior to analysis. Samples were stored at -20 °C until testing. The study was conducted in Bucharest, Romania, and all procedures used fully de-identified residual serum samples collected after the completion of routine diagnostic testing. Under the institutional and national regulatory standards, quality assurance evaluations using anonymized leftover specimens do not constitute interventional human subjects research. As such, the study was exempt from formal Ethics Committee approval and the requirement for informed consent, in accordance with the Declaration of Helsinki [14]. All procedures followed laboratory quality assurance guidelines set by the international standard ISO 15189:2022 [15], by the Clinical and Laboratory Standards Institute (CLSI) guidelines, including EP15-A3 (User Verification of Precision and Estimation of Bias) and EP07 (Interference Testing in Clinical Chemistry), under routine clinical laboratory conditions.

Table 1: The study protocol and materials used.

Study objectives					
Assay	**Precision**	**Analytical Sensitivity**	**Endogenous Interferents**	**Clinical Sensitivity**	**Clinical Specificity**
HIV 1/2 Combo	5 days (5 reps x 5 days) / Internal quality controls	1st WHO IRP HIV p24 Antigen panel (NIBSC 16/210)	Creatinine (15 mg/dL), Glucose (1000 mg/dL), Glycerol (18 mg/dL), Lactate (90 mg/dL), Pyruvate (5 mg/dL), Blood Urea Nitrogen (120 mg/dL), Uric Acid (25 mg/dL)	10 seroconversion panels and 15 subtype samples	1500 negative samples
Syphilis	5 days (5 reps x 5 days) / Internal quality controls	N/A		250 positives confirmed by RPR/VDRL/TPHA	250 negative samples

Abbreviations: WHO (World Health Organization); IRP (International Reference Panel); NIBSC (National Institute for Biological Standards and Control); N/A (Not applicable); RPR (Rapid Plasma Reagin); VDRL (Venereal Disease Research Laboratory); TPHA (Treponema Pallidum Hemagglutination Assay).

CLIA systems and immunoassays

Performance evaluations of the HIV 1/2 Ab-Ag Combo and Syphilis immunoassays (specific for each instrument) were conducted using three automated CLIA systems: the Abbott Alinity i (Abbott Laboratories Inc., Abbott Park, Illinois, USA), Snibe Maglumi X3 (Shenzhen New Industries Biomedical Engineering Co., Ltd., Shenzhen, China), and Mindray CL-1200i (Mindray Bio-Medical Electronics Co., Ltd., Shenzhen, China). While all three systems employ chemiluminescent detection principles, they differ in their luminescent labels, light emission kinetics, and technical specifications (Supplementary Table 1).

The evaluated HIV assays [16–18] are 4th generation chemiluminescent immunoassays designed for the simultaneous, qualitative detection of the HIV p24 antigen and antibodies to HIV-1 (group M and group O) and HIV-2. These assays serve as screening aids for HIV diagnosis and do not distinguish between antigen and antibody reactivity. All three device-specific assays share a similar analytical architecture: magnetic microparticles coated with HIV-1/HIV-2 antigens and anti-p24 antibodies capture the targets from the serum or plasma sample. Following a wash step, a labelled conjugate containing recombinant HIV-1/HIV-2 antigens and monoclonal anti-p24 antibodies is added. Subsequent washing and the addition of chemiluminescence catalysts generate a relative light unit (RLU) signal. This signal is directly proportional to the analyte concentration and is evaluated against a calibrated cutoff. The primary differences among the HIV assays stem from their specific conjugate labels, catalysts, and method traceability. The Maglumi X3 and CL-1200i assays are standardized against the 1st WHO International Reference Reagent for HIV-1 p24 Antigen (NIBSC code 90/636) [19]. Conversely, the Alinity i assay is referenced to the AFSSAPS international standard panel to maintain a consistent HIV-1 p24 antigen sensitivity of <50 pg/mL. These AFSSAPS standards were historically used to set lower limits of detection before widespread harmonization with the WHO international standard (IU/mL) [20]. Currently, no internationally accepted standards exist for anti-HIV 1 and anti-HIV 2 antibodies.

The evaluated syphilis assays [21–23] are qualitative CLIAs designed to detect total antibodies against Treponema pallidum (T. pallidum) as an aid in diagnosis. All three systems capture sample antibodies using microparticles coated with recombinant T. pallidum antigens, but they utilize different labelled conjugates to form the final sandwich complexes. The Alinity i and CL-1200i assays employ labelled anti-human IgM and IgG conjugates, whereas the Maglumi X3 utilizes a T. pallidum recombinant antigen conjugate. Like the HIV assays, the final RLU signal is proportional to the antibody concentration and compared to a calibrated cutoff. Because there are no international standards for syphilis testing, all three assays are referenced to distinct, manufacturer specific internal

standards, which contributes to the variations in their overall design and performance.

Analytical precision

Precision evaluation was conducted for all of the assays using the internal quality controls provided by each of the respective assay manufacturers, by performing 5 days of consecutive measurements, with 5 replicates in a single run on each day, with a total of 25 replicates for each quality control material. This protocol is consistent with the experimental design for precision verification as described in the Clinical and Laboratory Standards Institute (CLSI) EP15-A3 [24] guidelines (1 run x 5 reps x 5 days = 25 reps per level). The results for the precision study were compared between analyzers and relative to their respective manufacturer claimed performance.

Analytical sensitivity

Analytical sensitivity was evaluated for HIV 1/2 Combo Antigen-Antibody assays using the World Health Organization (WHO) 1st WHO Reference Panel for HIV-1 p24 Antigen (NIBSC code: 16/210) [25]. Panel members representing subtypes A1, B, C, D, G, H, circulating recombinant forms (CRFs): F1/CRF12_BF/BFrec, CRF20_BG, CRF01_AE, CRF02_AG, and group O, are provided lyophilized in individual vials. Panel member samples were reconstituted using 1 mL nuclease-free ultrapure water per vial on the day of testing and further used to prepare serial dilutions of the panel members. Unitage (IU/mL) values for the diluted samples were calculated according to the applied dilution factor, using the unitage values provided for the concentrated panel members. Although the WHO Expert Committee has not formally assigned a potency unitage to any of the panel members [26] the calibrated values relative to the 1st International Reference Reagent (NIBSC 90/636) are provided in International Units (IU) per 1 mL (IU/mL) for each panel member. Same day measurements were performed on all three analyzers by measuring 5 replicates for each dilution sample. Dilutions were further prepared and tested until two out of the three evaluated assays stopped detecting the tested dilution sample as reactive, returning negative results for all measurement replicates. Analytical sensitivity in this study was reported as a comparative performance characteristic and does not represent a formal limit of detection (LoD) determination. Because each dilution level was evaluated using 5 replicates rather than the larger replicate sizes required for statistical modeling, a formal probit analysis to calculate a 95% LoD was not feasible (see Statistical Analysis, below). Consequently, the reported sensitivity values should be interpreted strictly as comparative indicators of relative platform detection thresholds under identical experimental conditions, rather than absolute, established platform LoDs.

Endogenous interferences

Analytical interference for each assay was evaluated with seven common endogenous metabolites expected in a patient population using the ASSURANCE™ Interference Test Kit INT-06 (Sun Diagnostics, LLC, New Gloucester, Maine, USA) [27]. This kit contained the following metabolites, with targeted concentrations shown in parentheses: creatinine (15 mg/dL), glucose (1000 mg/dL), glycerol (18 mg/dL), lactate (90 mg/dL), pyruvate (5 mg/dL), blood urea nitrogen (120 mg/dL), uric acid (25 mg/dL). Leftover patient serum samples without any visible lipemia, haemolysis or icterus were used to prepare negative, low positive and high positive analyte pools. Serum pools were tested before spiking to establish the baseline measurand concentrations, then test pools were prepared by adding the interferent material to bring the test pool to the target interferent concentration. The required amount of interferent materials were calculated using the Assurance™ Excel Worksheets [27]. For each test pool, a matching control pool was prepared with the same volume of solvent (purified water) as used in preparation of the corresponding interferent pool. Same day measurements for each potential interferent were performed by measuring 6 replicates per run for each sample pool on all the analyzers. Furthermore, to minimize experimental error, the replicates were analyzed in alternating pattern between control pools and test pools. Because the evaluated immunoassays generate quantitative results that are interpreted qualitatively, interference acceptability was assessed relative to the assay cutoff rather than fixed percentage bias limits. Specifically, the maximum allowable difference (dmax) was established relative to the clinical decision point (immunoassay cutoff). For negative samples pool, dmax was defined as the maximum allowable signal shift that would not cause a true negative result to exceed the positivity cutoff. For the positives and the weak positives pools, dmax was defined as the maximum allowable signal shift that would not reduce a true positive result below the cutoff.

Clinical sensitivity

Clinical sensitivity evaluation was conducted differently for each assay (Table 1). For HIV 1/2 Combo assays, evaluation was conducted using one subtype performance panel (15 panel members of plasma samples collected worldwide from individual donors positive for HIV, consisting of 8 distinct HIV genotypes, including several circulating recombinant forms and mixed subtypes. Represented in this panel are B, BF, BG, CPX, C, CRF01_AE, CRF02_AG, and F1 subtypes) and 10 seroconversion panels, using the AccuSet HIV-1 Worldwide Performance Panel 0800-0492 [28] and the AccuVert HIV-1 Seroconversion Panels 0600-0243 [29]; 0600-0245 [30]; 0600-0248 [31]; 0600-0252 [32]; 0600-0258 [33]; 0600-0260 [34]; 0600-0262 [35]; 0600-0265 [36]; 0600-0271 [37]; 0600-0272 [38] (SeraCare, LGC Group, LGC Limited, Guildford, UK). Clinical sensitivity evaluation for syphilis was conducted using 250 leftover positive patient samples, confirmed positive by RPR/VDRL/TPHA testing. Because these specimens consisted of fully de-identified residual sera obtained after routine

diagnostic testing, matching clinical records regarding disease stage (primary, secondary, or latent), historical treatment status, or exact non-treponemal titers were completely inaccessible due to strict privacy and anonymization protocols.

Clinical specificity

Clinical specificity evaluation was conducted separately for each assay using leftover patient samples (Table 1). HIV 1/2 Combo assays evaluation was performed using 1500 unknown patient samples, while the evaluation for the Syphilis assays was performed using 250 diagnostic leftover patient samples. All leftover patient samples used for clinical specificity evaluation were selected from anonymised, previously tested negative samples, collected during daily routine from apparently healthy subjects from Romania, with no clinical indication of HIV or any other infectious disease symptoms. All reactive samples were tested in duplicate on the analyzer that provided the positive test result. Reactive results were further investigated by additional confirmatory testing methods (PCR, RPR/VDRL/TPHA), together with available clinical information, where applicable, to determine whether an active infection was present or to classify the result as false reactive.

Statistical analysis

Analyses were performed to evaluate the comparative performance of the three analytical systems. Precision was quantified using the coefficient of variation (CV%) and analyzed via analysis of variance (ANOVA, using the day of the run as random factor variable) to determine within run and total within laboratory performance. Regarding endogenous interference studies, patient serum samples were pooled to create standardized negative, low positive and high positive analyte levels. Pooling was necessary to provide sufficient sample volume for multi replicate testing and to establish a stable baseline measurand concentration before the addition of interferent materials. The acceptability of interference was assessed using the maximum allowable difference (dmax) relative to the medical decision point (assay cutoff).

Regarding sample size rationale, formal a priori power calculations were not performed because this study was designed as a non-interventional, post-market verification evaluation leveraging a prospective convenience cohort of residual clinical specimens. Instead, the sample sizes (1,500 negative patient samples for HIV specificity, 250 confirmed positive and 250 negative samples for syphilis, and 10 commercially sourced seroconversion panels of 132 individual specimens) were determined based on operational feasibility and the objective to meet or exceed international recommendations (CLSI and ISO 15189:2022) for laboratory immunoassay evaluations [15]. This study compared the three systems rather than formally establishing their primary baseline parameters. Given the global scarcity of commercial seroconversion panels, sample sizes were optimized for binomial precision; 1,500 HIV-negative samples ensure a 100% specificity estimate maintains a lower 95% confidence interval (CI) boundary of 99.75%, while 250 negative syphilis samples guarantee a lower 95% CI of 98.54%. This provides the high statistical precision needed to confidently isolate cross-platform performance differences.

For clinical sensitivity and specificity, point estimates were calculated as proportions of correctly identified samples, with 95% confidence intervals (CI) determined using the exact binomial Clopper-Pearson method. A formal statistical correction for multiple comparisons was not applied to the primary diagnostic metrics, as this study focused primarily on estimating diagnostic metrics via standalone point estimates and exact confidence intervals, rather than testing multiple simultaneous null hypotheses or relying on a large network of exploratory p-values. Pairwise cross-platform metrics are reported directly as descriptive and absolute clinical performance differences to ensure unadjusted data transparency for laboratory selection. To formally evaluate specific cross-platform diagnostic gaps, pairwise differences in clinical sensitivity between platforms were evaluated using the exact McNemar's test for paired nominal data, with a significance threshold established at $\alpha = 0.05$. Because the analytical sensitivity panel dilutions were evaluated using 5 replicates per concentration level, a formal Probit regression analysis to calculate a statistical 95% Limit of Detection (LoD) curve was not mathematically feasible. Consequently, these endpoint sensitivity values are strictly presented as operational, panel-specific detection thresholds rather than formalized statistical LoD metrics.

Finally, the robustness of each system against false reactivity was assessed by measuring the statistical separation between the negative sample distribution and the assay cutoff, expressed in standard deviation (SD). SD to cutoff ratio was calculated as the difference between the assay cutoff and the mean results signal, divided by the SD of the corresponding results distribution. A larger SD to cutoff value indicates greater separation from the cutoff and a lower likelihood of results misclassification due to analytical variation. This metric was used to assess how well the negative results were separated from the cutoff, thereby providing an estimate of assay robustness near cutoff and allowing comparison between assays independently of their reported signal units.

Additionally, the assessment of the reactive results distribution above assay cutoff assessment was included in order to determine the separation of the positive results relative to the cutoff, as low-positive results (gray-zone, borderline results) often fall within the analytical uncertainty range (more vulnerable to, and often caused by analytical interference). Thus, low positive results require a more cautious interpretive resolution than clearly reactive results located well above cutoff. The distribution of reactive results above the assay cutoff was plotted by grouping reactive sample results into value intervals relative to the cutoff (1.0-1.2, 1.2-1.4, 1.4-1.6, 1.6-1.8, 1.8-2.0, 2.0-3.0, 3.0-4.0, 4.0-5.0, and ≥5.0).

The proportion of reactive results within each interval was calculated using the total number of reactive results for that assay as the denominator.

Results

Precision evaluation results

The analytical precision of the three evaluated CLIA systems was determined by assessing both within-run repeatability and total within-laboratory precision across the evaluated assays (HIV 1/2 and Syphilis). Precision was quantified using the coefficient of variation (CV%) across replicates of multi-level control materials (Table 2A). All systems demonstrated acceptable overall precision profiles for routine diagnostic use. The obtained intermediate precision (CV%) was compared to the manufacturer's claimed performance (Table 2B).

For the HIV 1/2 Combo assays, Alinity i demonstrated excellent precision across all control levels (Table 2A), with measured precision (CV%) values below manufacturer claimed specifications (Table 2B). Maglumi X3 showed acceptable within-laboratory precision, with measured CV% values generally comparable to manufacturer claims, although higher than claimed variability was observed for several control levels. For CL-1200i, the measured CV% for the negative control (9.37%) exceeded the manufacturer claim (6.46%), it remained only marginally higher than the Alinity i (8.84%) and substantially lower than Maglumi X3 (17.32%), remaining within acceptable limits for qualitative infectious disease immunoassays. CL-1200i exhibited higher analytical precision than Maglumi X3 for both negative and positive antibody control materials (Table 2).

For the syphilis assays, all evaluated systems demonstrated acceptable within-laboratory precision under routine laboratory conditions (Table 2A). Measured precision (CV%) values were below manufacturer claimed limits for both control levels on Alinity i and Maglumi X3 (Table 2B). Measured CV% of the CL-1200i negative control (5.44%), although exceeding the manufacturer claim (3.29%), remained lower than the corresponding CV% observed for both Alinity i (11.07%) and Maglumi X3 (7.87%). Relatedly, the CV% of the CL-1200i for the positive control (2.30%) was only marginally higher than claimed (2.26%) and lower than observed variation for both Alinity i (2.83%) and Maglumi X3 (3.86%), While direct comparability is inherently limited by the use of distinct, manufacturer-specific control materials and varying concentrations, evaluating these overall trends indicates that the CL-1200i generally tended to exhibit the lowest overall imprecision across both control levels, whereas the Alinity i and Maglumi X3 demonstrated slightly higher variability, particularly near the negative baseline.

Table 2A: Precision evaluation results for HIV 1/2 Combo and Syphilis assays.

Assay	Material	Analyzer	N	Mean	Unit	Repeat. SD	Repeat. CV%	WithinLab SD	WithinLab CV%
HIV 1/2 Combo	Negative Control (-)	Alinity i	25	0.08	S/CO	0.005	6.00	0.007	8.84
		Maglumi X3	25	0.24	AU/mL	0.033	13.49	0.042	17.32
		CL1200i	25	0.1	COI	0.006	6.22	0.009	9.37
	Positive Control (1)	Alinity i	25	5.25	S/CO	0.068	1.30	0.068	1.30
		Maglumi X3	25	10.03	AU/mL	0.447	4.46	0.652	6.50
		CL1200i	25	4.86	COI	0.111	2.28	0.131	2.70
	Positive Control (2)	Alinity i	25	3.52	S/CO	0.067	1.90	0.078	2.21
		Maglumi X3	25	20.06	AU/mL	1.075	5.36	1.075	5.36
		CL1200i	25	8.11	COI	0.182	2.24	0.33	4.07
	Positive Control (3)	Alinity i	25	3.25	S/CO	0.044	1.37	0.054	1.66
		Maglumi X3	25	5.18	AU/mL	0.376	7.27	0.376	7.27
Syphilis	Negative Control (-)	Alinity i	25	0.04	S/CO	0.002	5.15	0.004	11.07
		Maglumi X3	25	0.31	AU/mL	0.023	7.52	0.024	7.87
		CL1200i	25	0.43	COI	0.010	2.46	0.023	5.44
	Positive Control (+)	Alinity i	25	2.55	S/CO	0.045	1.76	0.072	2.83
		Maglumi X3	25	10.28	AU/mL	0.350	3.40	0.40	3.86
		CL1200i	25	4.82	COI	0.109	2.27	0.111	2.30

Assay	Material	Analyzer	95% Lower SD	95% Lower CV%	95% Upper SD	95% Upper CV%
HIV 1/2 Combo	Negative Control (-)	Alinity i	0.005	6.12	0.013	15.86
		Maglumi X3	0.03	12.54	0.068	28.02
		CL1200i	0.006	6.45	0.016	17.09
	Positive Control (1)	Alinity i	0.053	1.02	0.095	1.81
		Maglumi X3	0.453	4.51	1.163	11.6
		CL1200i	0.098	2.01	0.200	4.11
	Positive Control (2)	Alinity i	0.058	1.66	0.117	3.33
		Maglumi X3	0.839	4.18	1.498	7.47
		CL1200i	0.218	2.68	0.681	8.39
	Positive Control (3)	Alinity i	0.040	1.23	0.084	2.59
		Maglumi X3	0.293	5.67	0.524	10.13
Syphilis	Negative Control (-)	Alinity i	0.003	7.09	0.009	24.83
		Maglumi X3	0.019	6.09	0.034	11.10
		CL1200i	0.015	3.47	0.053	12.36
	Positive Control (+)	Alinity i	0.049	1.92	0.138	5.39
		Maglumi X3	0.300	2.92	0.587	5.71
		CL1200i	0.086	1.79	0.155	3.21

Table 2B: Intermediate precision (Within-Lab CV%) compared to claimed values.

Assay	Control sample	N	Alinity i		Maglumi X3		CL1200i	
			Claimed	Measured	Claimed	Measured	Claimed	Measured
HIV 1/2 Combo	Negative Control	25	11.60%	8.84%	N/A	17.32%	6.46%	9.37%
	Positive Control 1	25	2.40%	1.30%	6.52%	4.46%	3.40%	2.70%
	Positive Control 2	25	5.00%	2.21%	4.65%	5.36%	3.66%	4.07%
	Positive Control 3	25	2.80%	1.66%	6.47%	N/A	N/A	N/A
Syphilis	Negative Control	25	12.10%	11.07%	N/A	7.87%	3.29%	5.44%
	Positive Control	25	4.40%	2.83%	4.29%	3.86%	2.26%	2.30%

Abbreviations: N/A (Not available). Green shading indicates CV% values meeting or exceeding manufacturer performance claims (≤ claimed CV%). Red shading indicates CV% values higher than the manufacturer's claimed precision. Yellow shading indicates CV% values marginally exceeding manufacturer claims.

Analytical sensitivity results for HIV 1/2 Combo

Analytical sensitivity results for the HIV 1/2 Combo assays for the 1st WHO HIV-1 p24 Ag panel (NIBSC 16/210) are summarized in Table 3. Alinity i HIV Combo Ab/Ag provided reactive results at every dilution tested for all subtypes/CRFs, including group O. Maglumi X3 remained positive down to 0.790 IU/mL for subtype B (panel member 16/214, dilution 1/10). CL-1200i returned positive results for antigen concentrations around 2.313 IU/mL, and reactive for group O at 2.05 IU/mL. For panel member 16/218 (subtype C) at dilution 1/15 corresponding to 0.767 IU/mL (not shown in Table 3), Maglumi X3 returned borderline negative results, very close to the assay cutoff (mean of 0.93 AU/mL for all 5 replicates, no result above assay cutoff), with reactivity close to cutoff (SD to cutoff 1.53, only 0.07 AU/mL below the cutoff). For panel member 16/222 (subtype F1/CRF12_BF/BFrec) at dilution 1/6 corresponding to a unitage of 2.12 IU/mL (not shown in Table 3), CL-1200i returned borderline negative results, very close to the assay cutoff (mean of 0.94 COI for all 5 replicates, no result above assay cutoff), indicating reactivity close to cutoff (SD to cutoff 1.74, only 0.06 COI below the cutoff). Analytical sensitivity evaluation revealed clear differences among the evaluated systems. Alinity i consistently demonstrated the highest analytical sensitivity, remaining reactive at the lowest tested antigen concentrations across all tested HIV-1 subtype samples. Maglumi X3 followed closely, demonstrating intermediate analytical sensitivity, while CL-1200i showed comparatively lower detection capability at low antigen concentrations.

Table 3: HIV-1 p24 Antigen analytical sensitivity of the HIV 1/2 Combo assays.

Panel members	Subtype	Unitage (IU/mL)	Alinity i	Maglumi X3	CL1200i
16/212	A1	8.8	Positive up to 1/12 (0.733 IU/mL)	Positive up to 1/10 (0.880 IU/mL)	Positive up to 1/4 (2.20 IU/mL)
16/214	B	7.9	Positive up to 1/12 (0.658 IU/mL)	Positive up to 1/10 (0.790 IU/mL)	Positive up to 1/3 (2.55 IU/mL)
16/216	B	8.4	Positive up to 1/11 (0.764 IU/mL)	Positive up to 1/10 (0.840 IU/mL)	Positive up to 1/3 (2.10 IU/mL)
16/218	C	11.5	Positive up to 1/16 (0.719 IU/mL)	Positive up to 1/14 (0.821 IU/mL)	Positive up to 1/5 (2.30 IU/mL)
16/220	D	9.0	Positive up to 1/12 (0.750 IU/mL)	Positive up to 1/11 (0.818 IU/mL)	Positive up to 1/4 (2.25 IU/mL)
16/222	F1/CRF12_BF/BFrec	12.7	Positive up to 1/18 (0.706 IU/mL)	Positive up to 1/16 (0.794 IU/mL)	Positive up to 1/5 (2.47 IU/mL)
16/224	G	8.0	Positive up to 1/11 (0.727 IU/mL)	Positive up to 1/10 (0.800 IU/mL)	Positive up to 1/3.5 (2.25 IU/mL)
16/226	CRF20_BG	9.9	Positive up to 1/14 (0.707 IU/mL)	Positive up to 1/12 (0.825 IU/mL)	Positive up to 1/4.5 (2.15 IU/mL)
16/228	CRF01_AE	10.3	Positive up to 1/14 (0.736 IU/mL)	Positive up to 1/13 (0.792 IU/mL)	Positive up to 1/4 (2.48 IU/mL)
16/230	CRF02_AG	4.0	Positive up to 1/5.5 (0.727 IU/mL)	Positive up to 1/5 (0.800 IU/mL)	Positive up to 1/1.5 (2.60 IU/mL)
16/232	H	6.0	Positive up to 1/8 (0.750 IU/mL)	Positive up to 1/7 (0.857 IU/mL)	Positive up to 1/2.5 (2.36 IU/mL)
16/234	group O	11.3	Positive up to 1/15 (0.753 IU/mL)	Positive up to 1/14 (0.807 IU/mL)	Positive up to 1/5.5 (2.05 IU/mL)
Mean unitage value (IU/mL) of antigen concentration for the last reactive dilutions across all panel members			**0.728 IU/mL**	**0.819 IU/mL**	**2.313 IU/mL**

Values in parentheses indicate the operational concentration of the final reactive dilution, calculated directly from the external WHO/NIBSC 16/210 panel member unitage and the applied dilution factor. Because these evaluated screeners are qualitative assays, these figures serve strictly as within-study relative endpoint reactivity comparisons under identical experimental conditions; they do not imply direct equivalence between the manufacturers' distinct master calibration frameworks (AFSSAPS mass-concentration for Alinity i vs. NIBSC 90/636 for Maglumi X3 and CL-1200i) nor do they represent formal platform LoD claims.

Potential endogenous interference evaluation results

For all the three CLIA systems, no clinically significant interference was observed from common endogenous substances tested according to CLSI EP07 (data not shown). Overall, Maglumi X3 immunoassays demonstrated slightly higher signal variation compared to Alinity i and CL-1200i. Although significant variations (in some cases exceeding ±10% relative difference between control and test pools) were observed in a few runs, the variation did not alter the qualitative interpretation of results (positive/negative). Notably, for HIV 1/2 Ab-Ag, Maglumi X3 showed concentration differences of 0.5-10.7% for creatinine (15 mg/dL), 0.8-14.9% for glucose (1000 mg/dL), 0.5-17.6% for glycerol (18 mg/dL), 3.3-13.7% for lactate (90 mg/dL) and 2.7-10.7% for uric acid (25 mg/dL) across negative, low positive, and positive pools, while CL-1200i showed differences of 3.7-15.1% for lactate (90 mg/dL). For all other endogenous metabolites, signal differences remained within ±10%. Accordingly, interference effects exceeding ±10% were not considered clinically significant as the qualitative results interpretation agreement between the control and test pools was maintained.

Clinical sensitivity results for HIV 1/2 Combo

Clinical sensitivity results for each individual HIV 1/2 seroconversion panel tested (10 panels), and the number of days since the 1st HIV positive duplicate test result is detected, are presented in Table 4. Across the HIV performance panel (data not shown), all 15 samples were identified as reactive by Alinity i and Maglumi X3. CL-1200i returned reactive results for 14/15 samples, with the exception of panel member 08 (0800-0492 [28]) - subtype C (data not shown), which was identified as reactive only by Alinity i and Maglumi X3 and not by CL-1200i.

Table 4: Days until 1st HIV detection (Seroconversion Panels).

No. Crt.	Seroconversion Panel ID	Alinity i	Maglumi X3	CL1200i
1	06000272	11	11	18
2	06000252	0	0	7
3	06000260	7	14	14
4	06000271	0	7	10
5	06000243	21	27	29
6	06000248	35	37	44
7	06000262	2	13	15
8	06000265	21	28	28
9	06000245	14	14	17
10	06000258	7	7	9
Mean number of days until detection		**11.8**	**15.8**	**19.1**
Number of positive results (out of 132 total specimens tested)		**78 (59.1%)**	**62 (46.9%)**	**46 (34.8%)**

Values shown for each seroconversion panel represent the day of the first reactive bleed detected by each assay. The mean number of days was calculated as the arithmetic mean of the day of first reactivity across all the 10 seroconversion panels for each assay. Number of positive results refers to the total number of reactive individual serial bleed panel member specimens detected by each assay out of 132 total specimens tested across all the seroconversion panels.

The clinical sensitivity evaluation revealed marked differences in early infection detection between the HIV-1/2 Combo assays (for panel-specific detection data, see Supplementary Table 2). Across the HIV panels (Table 4), Alinity i demonstrated the shortest average seroconversion period (mean of 11.8 days; median 9.0 with a range of 0-35), allowing HIV detection 7.3 days earlier compared to CL-1200i (mean of 19.1 days; median 16.0 with a range of 7-44), and 4 days earlier compared to Maglumi X3 (mean of 15.8 days; median 13.5 with a range of 0-37). Alinity i presented an increased average seroconversion sensitivity by 25% compared to CL-1200i, and by 12% compared to Maglumi X3, as Alinity i identified the highest proportion of early reactive results across seroconversion panels (59.1% positives detected). Maglumi X3 showed intermediate detection in selected panels (46.9% positives detected), while CL-1200i demonstrated lower average seroconversion sensitivity (34.8% positives detected) indicating a higher number of missed early infections relative to the other two assays. Although the assays differed in their overall mean time to detection and proportion of reactive specimens identified, several individual panel comparisons illustrate aligned performance during early HIV infection (p24 antigen detection).

The same 1st day of detection performance was observed for Alinity i and Maglumi X3 for four seroconversion panels: 0600-0272 (HIV-1 subtype B donor; 6 serial bleeds collected over 30 days); 0600-0252 (HIV-1 donor; 4 serial bleeds ranging from Fiebig Stage II to IV collected over 14 days); 0600-0258 (HIV-1 donor; 4 serial bleeds including Fiebig Stage I-II/III collected over 16 days); 0600-0245 (HIV-1 donor; 7 serial bleeds including Fiebig Stage I, II and II/III collected over 21 days). Notably, both Alinity i and Maglumi X3 detected early HIV infection: panel 0600-0252, starting with the 1st bleed (Day 0); panel 0600-0258, starting with 2nd bleed (Day 7); panel 0600-0245, starting with the 5th bleed (Day 14); all being equivalent to Fiebig Stage I (acute, early HIV infection) [39]. For Maglumi X3 and CL-1200i, the same 1st day of detection performance for early infection was observed for seroconversion panel 0600-0260 (HIV-1 donor; 5 serial bleeds including Fiebig Stage I to III collected over 14 days), starting with the 5th bleed (Day 14) equivalent to Fiebig Stage II (acute, early HIV infection) [39]. At these early stages (Fiebig I-II), infection is characterized by detectable plasma viral load HIV RNA PCR positive and p24 antigen positive (only by highly

sensitive 4th generation assays [40]), with negative HIV-1 antibody tests (as the immune system has not yet developed detectable antibodies) [41].

Clinical sensitivity results for syphilis

Clinical sensitivity results for the syphilis assays revealed comparable performance for Alinity i and Maglumi X3 (Table 5). Alinity i correctly identified 249 of 250 positive samples, corresponding to a clinical sensitivity of 99.60% (95% CI: 97.79-99.99%). Maglumi X3 correctly identified 248 of the 250 positive samples, corresponding to a clinical sensitivity of 99.20% (95% CI: 97.14-99.90%). In contrast, CL-1200i identified only 240 of 250 positive samples, obtaining a clinical sensitivity of 96.00% (95% CI: 92.77-98.07%), showing an increased rate of false negative results (10/250, 4%). Pairwise comparisons using the exact McNemar's test showed no statistically significant difference in sensitivity between Alinity i and Maglumi X3 (p = 1.000). However, both platforms demonstrated a statistically significant sensitivity advantage over the CL-1200i system (p = 0.012 for Alinity i vs. CL-1200i; p = 0.039 for Maglumi X3 vs. CL-1200i).

Table 5: Clinical sensitivity evaluation for syphilis assays.

Analyzer	Cutoff	Unit	N	Negative	Positive	Mean	Sensitivity	Sensitivity 95%CI
Alinity i	1	S/CO	250	1	249	13.93	99.60%	(97.79% / 99.99%)
Maglumi X3	1	AU/mL	250	2	248	85.85	99.20%	(97.14% / 99.90%)
CL1200i	1	COI	250	10	240	24.70	96.00%	(92.77% / 98.07%)

Although the majority of reactive samples on all systems were distributed well above the assay cutoff, the results distribution of results differed between assays (Supplementary Figure 1). The proportion of reactive results at ≥5.0 was 205/249 (82.3%) for Alinity i, 242/248 (97.6%) for Maglumi X3, and 220/240 (91.7%) for CL-1200i. Notably, Alinity i showed a higher proportion of low positive reactive results near the cutoff, with 15/249 (6%) reactive sample results falling between the value range of 1.2 and 2.2, compared with 3/248 (1.2%) for Maglumi X3 and 5/240 (2.1%) falling between the 1.0 and 2.2 value range for CL-1200i (Supplementary Figure 1). Maglumi X3 showed the greatest separation of reactive results from cutoff, with the highest proportion of strongly reactive results at ≥5.0 (242/248, 97.6%) and the lowest proportion of low-positive reactive results near cutoff in the 1.2-2.2 range (3/248, 1.2%). Alinity i showed the opposite pattern, with the lowest proportion of results above ≥5.0 (82.3%) and the highest proportion of reactive results clustered near cutoff (6%), while CL-1200i showed an intermediate pattern (91.7% results ≥5; 2.1% results between 1.0-2.2). Despite showing a greater proportion of reactive results clustered close to the cutoff, Alinity i demonstrated the highest clinical sensitivity, achieving reliable detection of the positive samples.

Clinical specificity evaluation results

Clinical specificity results are summarized in Table 6. For the HIV 1/2 Combo assays, from a total of 1,500 leftover patient samples tested, 12 samples were found reactive (returning duplicate positive results). All reactive samples were confirmed negative by confirmatory HIV-1/HIV-2 RNA PCR testing and were considered false positive results. Alinity i demonstrated a clinical specificity of 99.87% (95% CI: 99.52-99.98%), with 2 false reactive results (0.13% false positive rate), followed closely by Maglumi X3 and CL-1200i, both showed a specificity of 99.67% (95% CI: 99.22-99.89%), corresponding to 5 false reactive results for each system (0.33% false positive rate). The majority of negative samples across all HIV assays clustered well below the respective assay cutoffs (SD to cutoff 13.55 for Alinity i, 5.70 for Maglumi X3, 6.84 for CL-1200i) (Supplementary Figure 2A).

For syphilis, from a total of 250 leftover patient samples tested, 9 samples were found reactive on at least one of the analyzers. From these, two samples were found positive on two analyzers (one sample returned duplicate positive results on both Alinity i and CL-1200i, and another sample returned duplicate positive results on both Maglumi X3 and CL-1200i). All the reactive samples were confirmed negative by additional confirmatory testing (RPR/VDRL/TPHA) and were classified as false positive results. Both Alinity i and Maglumi X3 showed a specificity of 99.20% (95% CI: 97.14-99.90%), corresponding to 2 false reactive results (0.80% false positive rate). CL-1200i demonstrated a specificity of 97.20% (95% CI: 94.32-98.87%), corresponding to 7 false reactive results out of the 250 negative tested samples (2.80% false positive rate). The majority of negative samples across all assays clustered well below the respective assay cutoffs (SD to cutoff 4.15 for Alinity i, 2.35 for Maglumi X3, 1.47 for CL-1200i) (Supplementary Figure 2B).

Table 6: Clinical specificity for HIV 1/2 and Syphilis.

Assay	Analyzer	Cutoff	Unit	N	Negative	Positive	Mean	Specificity	Specificity 95%CI
HIV 1/2 Combo	Alinity i	1	S/CO	1500	1498	2	0.13	99.87%	(99.52% / 99.98%)
	Maglumi X3	1	AU/mL	1500	1495	5	0.21	99.67%	(99.22% / 99.89%)
	CL1200i	1	COI	1500	1495	5	0.12	99.67%	(99.22% / 99.89%)
Syphilis	Alinity i	1	S/CO	250	248	2	0.08	99.20%	(97.14% / 99.90%)
	Maglumi X3	1	AU/mL	250	248	2	0.33	99.20%	(97.14% / 99.90%)
	CL1200i	1	COI	250	243	7	0.48	97.20%	(94.32% / 98.87%)

Discussion

The rapid advancement and widespread implementation of CLIA platforms represents a critical milestone in the management of infectious diseases. However, the success of global public health strategies relies heavily on the real-world performance of these diagnostic tools. The need for accurate serological screening cannot be overstated, particularly given that syphilis continues to re-emerge globally and has been consistently associated with an increased risk of HIV acquisition [42]. Achieving the WHO goals to eliminate HIV, viral hepatitis, and sexually transmitted infections as public health threats by 2030 requires highly sensitive diagnostic frameworks. For instance, maternal syphilis, if untreated or treated too late, results in congenital infection in 50% to 80% of cases alongside severe adverse birth outcomes [9]. Similarly, early detection of HIV infection remains critical for blood safety and global elimination efforts. To address these diagnostic challenges, our study provides a comprehensive performance evaluation of the Alinity i, Maglumi X3, and CL-1200i systems under routine clinical laboratory conditions.

It is important to note that all three CLIA systems are deployed for high-throughput screening, they do not share identical cleared clinical applications. The Alinity i syphilis assay is explicitly indicated as a primary screening test to prevent infectious transmission via blood and tissues, whereas the manufacturers of the CL-1200i and Maglumi X3 systems explicitly exclude their use for evaluating blood donation or transplantation suitability [16,22,23]. These structural parameters offer an essential baseline context when analyzing the variations in performance across the systems.

Our findings revealed that all tested CLIA platforms exhibit acceptable overall precision profiles for routine diagnostic use. Instances in which manufacturer claimed precision specifications were not fully met likely reflect differences between controlled manufacturer validation environments and routine clinical laboratory conditions rather than clinically meaningful analytical limitations. For example, relatively higher variation observed at the negative control level remained well below assay cutoff thresholds and did not impact clinical interpretation. Furthermore, endogenous interference testing confirmed the robustness of the tested CLIA systems against common biochemical interferents that could be found in patients with metabolic disturbances or renal dysfunction.

Analytical sensitivity in HIV screening is primarily driven by the ability to detect low-level p24 antigen during the early phase of infection, prior to the development of a fully established antibody response. Evaluating assay performance using progressive dilutions of diverse, recombinant HIV panels is critical because it rigorously simulates the low biomarker concentrations characteristic of acute or suppressed infections across a globally representative array of viral subtypes [11,43]. Because these pharmacological interventions can profoundly blunt p24 antigenemia and delay antibody maturation, individuals experiencing breakthrough infections often present with atypical, highly suppressed serological profiles. Earlier detection during seroconversion is particularly important for blood safety screening, as p24 antigen may be the only detectable marker before antibody maturation.

Our evaluation revealed clear performance differences, with the Alinity i platform allowing on average HIV detection at least 7.3 days earlier compared to CL-1200i, and at least 4 days earlier compared to Maglumi X3. This cross-platform performance hierarchy was fully supported by median values (Alinity i: 9.0 days; Maglumi X3: 13.5 days; CL-1200i: 16.0 days). Furthermore, a leave-one-out sensitivity analysis (data not shown) confirmed that no single panel disproportionately drove these observed variations, with baseline mean differences shifting by less than 1.0 day across all iterations. Such differences are likely attributable to variations in assay design and the specific biological reagents employed (e.g. recombinant antigens and monoclonal antibodies) alongside differences in signal amplification chemistry and detection kinetics among the systems. When interpreting these variations, the underlying metrological non-equivalence of the assays' baseline calibration standards must be considered. The Alinity i system relies on traceability to AFSSAPS mass-concentration definitions (<50 pg/mL of p24), whereas the Maglumi X3 and CL-1200i platforms are calibrated against NIBSC 90/636 international activity units (IU/mL). While utilizing a standardized external reference panel (WHO/NIBSC 16/210) provides an operational look at relative endpoint reactivity under identical conditions, these distinct manufacturing calibration frameworks prevent direct numerical translation and reinforce why these metrics reflect panel-specific performance rather than interchangeable absolute sensitivities.

Ultimately, these variations highlight key performance

characteristics with potential clinical implications for early infection identification. Maximized analytical sensitivity across diverse viral subtypes is essential not only for monitoring specific clinical cohorts but for routine diagnostic testing across all settings. The ability to reliably detect low-level biomarkers during the earliest stages of infection is vital for guiding prompt treatment initiation, optimizing patient outcomes, and implementing effective public health strategies to prevent ongoing HIV transmission. This clinical significance is further clarified by valuating performance against the panels' manufacturer-verified NAT data. Tracking the residual diagnostic window (i.e. the days between initial RNA detection and immunoassay reactivity) reveals clear cross-platform gaps. High-sensitivity systems often closed the serological window immediately (0 days) or within 4 days of NAT positivity, whereas less sensitive architectures left extended diagnostic gaps of up to 7 to 11 days.

The clinical relevance of sensitivity differences must also be interpreted in the context of the global re-emergence of syphilis. In such settings, even modest reductions in assay sensitivity may translate into a meaningful number of missed infectious cases, with implications for transmission control, prenatal screening and congenital syphilis prevention. From a clinical and public health perspective, failure to detect 4% of confirmed positive samples represents a meaningful diagnostic gap for the CL-1200i platform. This performance tiering was statistically verified by the exact McNemar's test; while Alinity i and Maglumi X3 share statistical equivalence, both platforms achieved a statistically superior diagnostic threshold over the CL-1200i architecture.

The observed differences underscore the importance of assay selection in clinical contexts where early detection is critical. Indeed, as mentioned previously, this analytical divergence directly aligns with the CLIA systems' underlying intended use scopes [16,22,23]. Additionally, because syphilis CLIA immunoassays can be used in reverse diagnostic algorithms, discordant screening results necessitate downstream non-treponemal and treponemal confirmatory testing to accurately differentiate active infections. Specifically, the intended use statements for Maglumi X3 warn about these limitations [22]. Clinical specificity was highly preserved across all platforms, though minor variations were observed. In high volume screening environments, even small differences in false positive rates may have operational implications, including increased confirmatory testing, and additional laboratory workload. Together, our results show that post-market evaluations can be very useful in revealing real-world performance differences that may not be fully captured during manufacturer in-house validation, particularly under the regulatory framework established by the EU IVDR [44]. Our study provides a robust and comprehensive evaluation of infectious disease immunoassays under authentic clinical laboratory conditions. By utilizing highly characterized WHO reference materials, standardized seroconversion panels, and a large cohort of clinical diagnostic samples, our assessment yields highly reliable and actionable insights into the real-world performance of the CLIA systems. These assays offer same-day results aiding in diagnosis more swiftly than western blotting (HIV) and RPR/VDRL/TPHA (syphilis), thereby reducing the percentage of loss to follow-up [45,46]. For HIV, WHO testing strategies recommend national algorithms that use immunoassays instead of nucleic acid testing techniques and western blotting for the diagnosis of subjects over 18 months of age [45]. Similarly, for syphilis, the high-throughput capabilities of automated CLIAs have driven a global shift toward the "reverse" screening algorithm. By utilizing a highly sensitive treponemal immunoassay as the initial screening step rather than a traditional non-treponemal test like RPR, the reverse algorithm significantly improves the detection of early primary and late latent infections while reducing overall laboratory costs and turnaround times [46,47]. Ultimately, improving throughput capacity and time-to-result across both targets facilitates rapid linkage to prevention and treatment services, directly supporting the WHO goal of ending sexually transmitted infections as public health threats by 2030[3].

Nevertheless, certain limitations must be acknowledged. The generalizability of our findings is bounded by the study's single-center design utilizing a regional Romanian patient cohort, alongside the omission of specific manufacturing reagent lot numbers and instrument software versions. While these omissions restrict absolute longitudinal replication across diverse global demographics, the single-site framework successfully isolated the platforms from inter-laboratory environmental noise. Furthermore, genotype diversity was inherently limited to the available commercial panels and reference materials. Similarly, the lack of clinical staging, historical treatment documentation, and individual reference titer values for the syphilis-positive cohort prevents a stage-specific sensitivity analysis or a definitive clinical root-cause determination for the observed assay failures. Additionally, a limitation of the endogenous interference assessment was the utilization of pooled patient sera rather than individual clinical samples. While this pooling approach was operationally mandatory to secure the high sample volumes (~20 mL per interferent) required for parallel testing across three automated platforms, and directly complied with the commercial kit's standardized protocol for comparing baseline and spiked matrices, it inherently minimizes individual patient matrix variability near the assay cutoff thresholds. Lastly, parameters such as long-term reagent stability and lot to lot variability were not explicitly evaluated. Consequently, the reported performance characteristics strictly reflect the specific testing conditions present during the study period and may exhibit temporal or geographic variations.

Despite these constraints, the foundational evidence presented here can provide valuable guidance for CLIA platform selection and optimizing blood safety protocols. While all evaluated systems demonstrated acceptable precision and robustness for

routine laboratory use, their diagnostic sensitivity capabilities varied considerably. Specifically, the Alinity i system provided the earliest detection of HIV seroconversion, whereas the CL-1200i platform exhibited a notable diagnostic gap in clinical syphilis sensitivity. Ultimately, aligning system capabilities with strict clinical intended uses is essential to support global public health initiatives and infectious disease elimination targets. Future research that incorporates multicenter evaluations and longitudinal stability assessments have the potential to provide more information and extend the findings of the comparative assessment presented herein. Furthermore, studies aimed at targeted prospective evaluations within PrEP-using and early-ART cohorts may provide important information on how these analytical variations impact patient outcomes. Ultimately, continuous real-world validation frameworks are essential for maintaining diagnostic excellence and complying with stringent regulatory requirements.

Conflict of interest statement

D.G.C., S.D., and A.J.N. are affiliated to Dig Worldwide Ltd. Dig Worldwide Ltd was involved in the study design, data collection, analysis, and the decision to submit the manuscript for publication. D.G.C. is a contracted specialist working with Dig Worldwide Ltd, S.D. is an employee of, and A.J.N. is a shareholder of Dig Worldwide Ltd. Abbott Laboratories had visibility towards the data and the manuscript before submission without altering the results or their interpretation. The final publication decision remained independently with the authors. The authors declare no other competing interests.

Funding statement

This study was supported by Dig Worldwide Ltd., a commercial consultancy firm in the life sciences sector, utilizing project funding from Abbott Laboratories. This research was an initiative of Dig Worldwide Ltd. and the results presented in this study were not influenced or biased in any way by CLIA system manufacturers or other second or third parties.

Data availability statement

The analyzed data that support the findings of this study are presented in the article and included in the supplementary materials, detailed datasets and further information inquiries are available on request from the corresponding author, by emailing: dgc@dig-worldwide.com.

CRediT authorship contribution statement

Author Contributions: Conceptualization, D.C., S.D., A.N.; validation D.C., S.D., A.N.; investigation D.C., formal analysis D.C., S.D., A.N.; data curation D.C.; writing - original draft D.C.; writing - review & editing D.C., S.D., A.N.; visualization D.C., project administration D.C., S.D., A.N.; supervision A.N.; funding acquisition A.N.

References

1. World Health Organization. Global health sector strategies on HIV, viral hepatitis and sexually transmitted infections for the period 2022-2030 https://www.who.int/publications/i/item/9789240053779 (accessed April 2, 2026).
2. CDC Reports Latest National Data on Syphilis in Newborns and Sexually Transmitted Infections (STIs) | CDC Newsroom n.d. https://www.cdc.gov/media/releases/2025/2025-cdc-reports-latest-national-data-on-syphilis-in-newborns-and-sexually-transmitted-infections-stis.html (accessed April 2, 2026).
3. World Health Organization. Global health sector strategies on, respectively, HIV, viral hepatitis and sexually transmitted infections for the period 2022-2030. GHSS 2022-2030 2022.
4. UNAIDS 2025 Global AIDS Update — AIDS, Crisis and the Power to Transform | UNAIDS n.d. https://www.unaids.org/en/resources/documents/2025/2025-global-aids-update-summary (accessed April 2, 2026).
5. Ghanem KG, Ram S, Rice PA. The Modern Epidemic of Syphilis. New England Journal of Medicine 2020;382:845-854. https://doi.org/10.1056/NEJMRA1901593;WGROUP:STRING:MMS.
6. WHO. Global Health Sector Strategies on HIV, Viral Hepatitis and the sexually transmitted infections. World Health Organization 2021;ector stra:2022-2230.
7. Kamb ML, Newman LM, Riley PL, Mark J, Hawkes SJ, Malik T, et al. A Road Map for the Global Elimination of Congenital Syphilis. Obstet Gynecol Int 2010;2010:312798. https://doi.org/10.1155/2010/312798.
8. U.S. Syphilis Cases in Newborns Continue to Increase: A 10-Times Increase Over a Decade | CDC Online Newsroom | CDC n.d. https://www.cdc.gov/media/releases/2023/s1107-newborn-syphilis.html (accessed April 2, 2026).
9. CDC Reports Latest National Data on Syphilis in Newborns and Sexually Transmitted Infections (STIs) | CDC Newsroom n.d. https://www.cdc.gov/media/releases/2025/2025-cdc-reports-latest-national-data-on-syphilis-in-newborns-and-sexually-transmitted-infections-stis.html (accessed April 2, 2026).
10. Moschese D, Lazzarin S, Colombo ML, Caruso F, Giacomelli A, Antinori S, et al. Breakthrough Acute HIV Infections among Pre-Exposure Prophylaxis Users with High Adherence: A Narrative Review. Viruses 2024, Vol 16, 2024;16. https://doi.org/10.3390/V16060951.
11. Seed CR, Styles CE, Hoad VC, Yang H, Thomas MJ, Gosbell IB. Effect of HIV pre-exposure prophylaxis (PrEP) on detection of early infection and its impact on the appropriate post-PrEP deferral period. Vox Sang 2021;116:379-387. https://doi.org/10.1111/VOX.13011.
12. Rodgers MA, Vallari AS, Yamaguchi J, Holzmayer V, Harris B, Toure-Kane C, et al. ARCHITECT HIV Combo

Ag/Ab and RealTime HIV-1 Assays Detect Diverse HIV Strains in Clinical Specimens. AIDS Res Hum Retroviruses 2018;34:314. https://doi.org/10.1089/AID.2017.0244.
13. Miedouge M, Grèze M, Bailly A, Izopet J. Analytical sensitivity of four HIV combined antigen/antibody assays using the p24 WHO standard. Journal of Clinical Virology 2011;50:57–60. https://doi.org/10.1016/j.jcv.2010.09.003.
14. Association WM. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA 2013;310:2191-2194. https://doi.org/10.1001/JAMA.2013.281053.
15. International Organization for Standardization (ISO). ISO 15189:2022. Medical laboratories – requirements for quality and competence 2022.
16. Abbott Laboratories Abbott Ireland Diagnostics Division. Alinity i HIV Ag/Ab Combo Reagent Kit, Ref. No. 08P07, Instructions for Use [Package Insert]
17. Mindray Bio-Medical Electronics Co. Ltd. CL-series HIV Antigen and Antibodies to Human Immunodeficiency Virus (CLIA), Ref. No. HIV121, Instructions for Use [Package Insert]
18. Shenzhen New Industries Biomedical Engineering Co. Ltd. MAGLUMI® HIV Ab/Ag Combi (CLIA), Ref. No. 130219008M, Instructions for Use [Package Insert]
19. NIBSC (National Institute for Biological Standards and Control), WHO International Laboratory for Biological Standards, UK Official Medicines Control Laboratory. International Reference Reagent, HIV-1 P24 Antigen, NIBS code: 90/636, Instructions for Use [Package Insert Version 6.0, Dated 12/12/2023].
20. Miedouge M, Grèze M, Bailly A, Izopet J. Analytical sensitivity of four HIV combined antigen/antibody assays using the p24 WHO standard. Journal of Clinical Virology 2011;50:57–60. https://doi.org/10.1016/j.jcv.2010.09.003.
21. Abbott Laboratories Abbott Ireland Diagnostics Division. Alinity i Syphilis Reagent Kit, Ref. No. 07P60, Instructions for Use [Package Insert]
22. Shenzhen New Industries Biomedical Engineering Co. Ltd. MAGLUMI® Syphilis (CLIA), Ref. No. 130269003M, Instructions for Use [Package Insert]
23. Mindray Bio-Medical Electronics Co. Ltd. CL-series Anti-TP, Antibody to Treponema Pallidum (CLIA), Ref. No. 105-005671-00, Instructions for Use [Package Insert]
24. Clinical and Laboratory Standards Institute. CLSI EP15-A3. User Verification of Precision and Estimation of Bias; Approved Guideline - 3rd Edition 2014.
25. NIBSC (National Institute for Biological Standards and Control), WHO International Laboratory for Biological Standards, UK Official Medicines Control Laboratory. WHO Reference Panel, 1st WHO International Reference Panel for HIV-1 p24 Antigen, NIBS code: 16/210, Instructions for Use [Package Insert Version 3.0, Dated 23/09/2024]
26. World Health Organization. WHO Expert Committee on Biological Standardization - Sixty-ninth report, Section 7.1.2 First WHO International Reference Panel for HIV-1 p24 antigen and First WHO International Reference Reagent for HIV-2 p26 antigen. WHO Technical Report Series, No 1016, 2019
27. LLC SD. ASSURANCETM Interference Test Kit INT-06 Instructions for Use [Package Insert]; Sample Interference Testing procedure (based on CLSI EP-7A guidelines), dilution calculator and results reporting spreadsheet
28. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuSetTM HIV-1 Worldwide Performance Panel, Ref. 0800-0492, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuSet-HIV-Worldwide-Performance-Panel-0800-0492/ 2021.
29. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0243, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000243-0600-0243/ 2023.
30. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0245, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000245-0600-0245/ 2023.
31. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0248, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000248-0600-0248/ 2022.
32. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0252, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000252-0600-0252/ 2022.
33. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0258, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000258-0600-0258/ 2024.
34. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0260, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000260-0600-0260/ 2023.
35. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0262, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000262-0600-0262/ 2023.
36. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0265, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000265-0600-0265/ 2021.

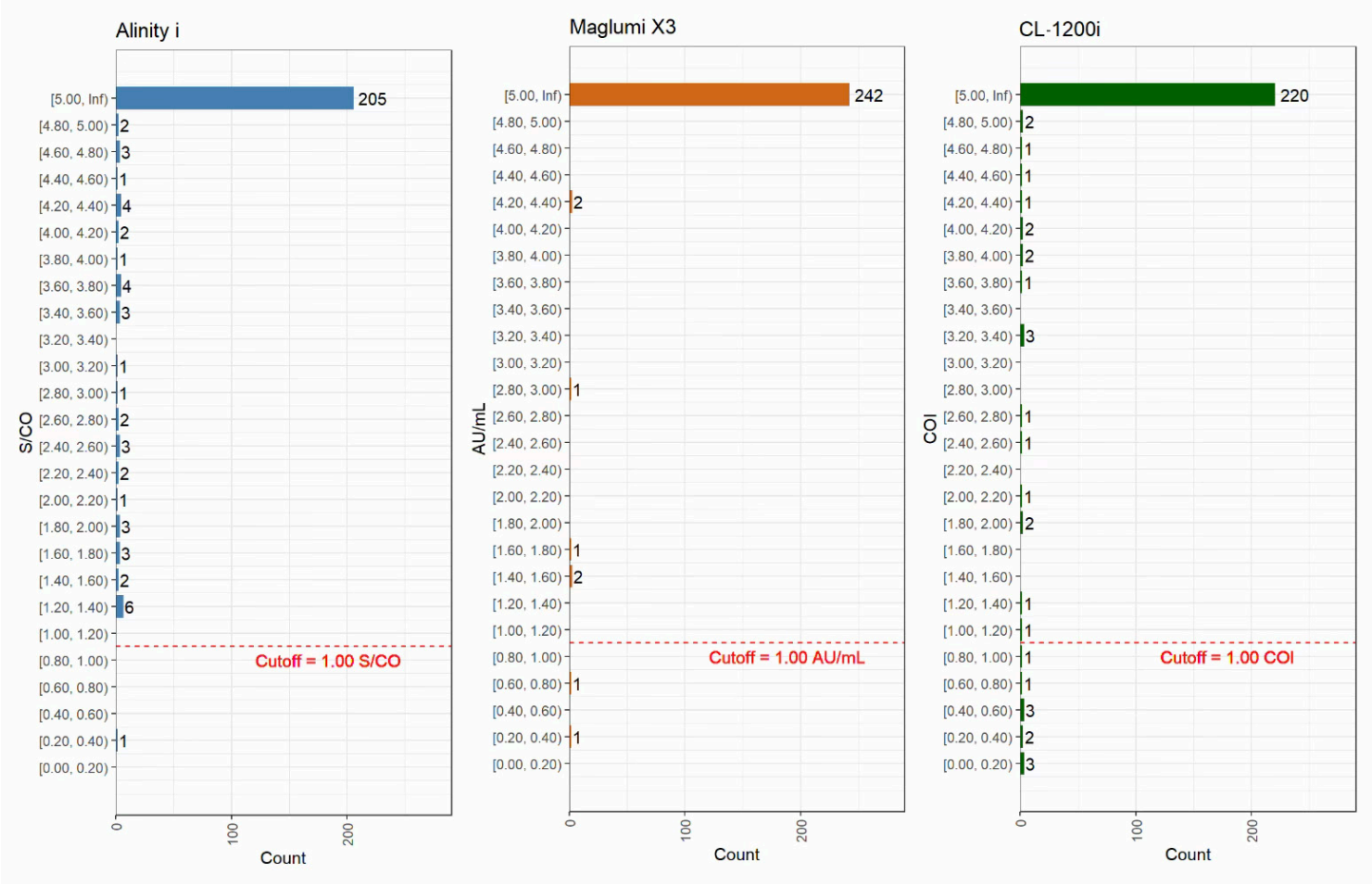
37. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0271, Instructions for Use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000271-0600-0271/ 2024.
38. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0272, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000272-0600-0272/ 2022.
39. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 2003;17:1871-1879. https://doi.org/10.1097/00002030-200309050-00005.
40. Pandori MW, Hackett J, Louie B, Vallari A, Dowling T, Liska S, et al. Assessment of the Ability of a Fourth-Generation Immunoassay for Human Immunodeficiency Virus (HIV) Antibody and p24 Antigen To Detect both Acute and Recent HIV Infections in a High-Risk Setting. J Clin Microbiol 2009;47:2639. https://doi.org/10.1128/JCM.00119-09.
41. Chéret A. Acute HIV-1 Infection: Paradigm and Singularity. Viruses 2025;17:366. https://doi.org/10.3390/v17030366.
42. Gaspari V, Filippini A, Orioni G, Mussi M, Carpanese MA, La Placa M, et al. The Rise of Syphilis Infections and Reinfections over a Decade (2009–2019) in the Bolognese Area: A Retrospective Analysis. Microorganisms 2025;13:285. https://doi.org/10.3390/microorganisms13020285.
43. Nair M, Gettins L, Fuller M, Kirtley S, Hemelaar J. Global and regional genetic diversity of HIV-1 in 2010–21: systematic review and analysis of prevalence. Lancet Microbe 2024;5. https://doi.org/10.1016/S2666-5247(24)00151-4.
44. Regulation - 2017/746 - EN - Medical Device Regulation - EUR-Lex n.d. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32017R0746 (accessed April 2, 2026).
45. World Health Organization (Geneva). World Health Organization (2024). Consolidated guidelines on differentiated HIV testing services n.d.
46. Totten YR, Hardy BM, Bennett B, Rowlinson MC, Crowe S. Comparative Performance of the Reverse Algorithm Using Architect Syphilis TP Versus the Traditional Algorithm Using Rapid Plasma Reagin in Florida's Public Health Testing Population. Ann Lab Med 2019;39:396-399. https://doi.org/10.3343/ALM.2019.39.4.396.
47. Morshed MG, Singhb AE. Recent trends in the serologic diagnosis of syphilis. Clin Vaccine Immunol 2015;22:137-143. https://doi.org/10.1128/CVI.00681-14.

Supplementary Table 1: Characteristics and technical specifications of each system.

	Alinity i	Maglumi X3	CL1200i
Manufacturer	Abbott Laboratories Inc.	Shenzhen New Industries Biomedical Engineering (Snibe) Co., Ltd.	Mindray BioMedical Electronics Co., Ltd.
Analyzer model	Alinity i	Maglumi X3	CL1200i
Method	Chemiluminescent microparticle immunoassay (CMIA)	Chemiluminescent microparticle immunoassay (CMIA)	Chemiluminescent microparticle immunoassay (CMIA)
Measurement principle	Flash CLIA	Flash CLIA	Glow CLIA
CLIA Label	Acridinium	ABEI (Isoluminol derivate)	ALP
CLIA Substrate	NaOH, H2O2	NaOH, H2O2	AMPPD
Solid phase	paramagnetic microparticles	paramagnetic nanoparticles	paramagnetic microparticles
Capacity	200 tests/hour	200 tests/hour	180 tests/hour

CLIA (chemiluminescent immunoassay), ABEI [N-(4-Aminobutyl)-N-ethylisoluminol], NaOH (sodium hydroxide), H2O2 (hydrogen peroxide), AMPPD [3-(2`-spiroadamantane)-4-methoxy-4-(3``-phosphoryloxy)phenyl-1, 2-dioxietane disodium salt], ALP (Alkaline phosphatase).

Supplementary Figure 1: Distribution of the results obtained for the clinical sensitivity evaluation of the syphilis antibody assays on the three automated CLIA platforms (Alinity i, Maglumi X3, and CL-1200i).



The figure summarizes results for 250 residual serum samples confirmed positive for syphilis by reference serology (RPR/VDRL/TPHA). In each panel, horizontal bars indicate the number of samples within each signal values interval, and the dashed red line indicates the reactive assay cutoff. Results are reported in assay specific units (S/CO for Alinity i, AU/mL for Maglumi X3, and COI for CL-1200i); therefore, the distributions should be interpreted mainly by their separation from the cutoff within each platform, rather than by direct comparison of absolute numeric values across analyzers. Most positive samples were strongly reactive, with results ≥5.0 observed in 205/249 reactive samples (82.3%) on Alinity i, 242/248 (97.6%) on Maglumi X3, and 220/240 (91.7%) on CL-1200i. Low-positive results near the cutoff were more frequent on Alinity i, with 15/249 reactive samples (6.0%) between 1.2 and 2.2, compared with 3/248 (1.2%) on Maglumi X3 and 5/240 (2.1%) on CL-1200i. Samples falling below the cutoff represent false-negative results detected on each assay (1/250 on Alinity i, 2/250 on Maglumi X3, and 10/250 on CL-1200i), respectively, corresponding to clinical sensitivities of 99.60% (Alinity i), 99.20% (Maglumi X3) and 96.00% (CL-1200i). Specifically, the 10 false-negative specimens on the CL-1200i platform were distributed across the following non-reactive value ranges: 3 specimens within [0.00, 0.20) COI, 3 specimens within [0.20, 0.40) COI, 3 specimens within [0.40, 0.60) COI, and 1 specimen within [0.80, 1.00) COI. The profound suppression of signal (≤0.60 COI) in 90% of these missed cases indicates a structural absence of antibody recognition rather than borderline imprecision, likely corresponding to early primary infections or historical, successfully treated latent cases with low residual antibody titers. Overall, the figure shows that Maglumi X3 produced the tightest clustering of strongly reactive results, whereas Alinity i, despite a larger fraction of near cutoff positives, achieved the highest overall sensitivity.

Supplementary Figure 2: Distribution of the results obtained for the clinical specificity evaluation of the HIV 1/2 Combo assays (A) and syphilis assays (B) on the three automated CLIA platforms (left to right: Alinity i, Maglumi X3, and CL-1200i).



Horizontal bars show the number of results within each interval; the dashed red line marks the reactive assay cutoff. Panel (A) summarizes results for 1,500 HIV negative residual samples, corresponding to clinical specificities of 99.87% for Alinity i (2/1500 false positives), 99.67% for Maglumi X3 (5/1500), and 99.67% for CL-1200i (5/1500). The SD to cutoff (13.55 for Alinity i, 5.70 for Maglumi X3, and 6.84 for CL-1200i) indicating the greatest negative result separation on Alinity i. Panel (B) summarizes results for 250 syphilis negative residual samples. Specificity was 99.20% for Alinity i (2/250 false-positives), 99.20% for Maglumi X3 (2/250) and 97.20% for CL-1200i (7/250); the corresponding SD to cutoff values were 4.15, 2.35, and 1.47, respectively. Samples above cutoff represent false reactive results (resolved by confirmatory testing). This figure therefore illustrates not just overall specificity, but also how the negative results are separated from the reactive cutoff threshold.

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

An Exploratory Study to assess the Feasibility and Acceptability of Clinical Infection Score Estimation in Patients with Suspected Infection at a tertiary care hospital

Prashanth Mada[1], Radhika Soanker[2*], Sakthivadivel Varatharajan[3], Shyamala Ravikoti[4], Shantveer G Uppin[5], Tara Roshni Paul[5], Gaddala Mary Sowjanya[6]

[1]*Department of Forensic Medicine and Toxicology, All India Institute of Medical Sciences, Bibinagar, Hyderabad Metropolitan Region, India*

[2]*Department of Pharmacology, All India Institute of Medical Sciences, Bibinagar, Hyderabad Metropolitan Region, India*

[3]*Department of General Medicine, Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry, India*

[4]*Department of Microbiology, All India Institute of Medical Sciences, Bibinagar, Hyderabad Metropolitan Region, India*

[5]*Department of Pathology, Nizam's Institute of Medical Sciences, Hyderabad, India*

[6]*Department of Forensic Medicine and Toxicology, Kamineni Academy of Medical Sciences and Research Centre, Hyderabad, India*

Article Info

Corresponding Author:
Radhika Soanker
Department of Pharmacology,
All India Institute of Medical Sciences, Bibinagar,
Hyderabad Metropolitan Region, India
E-mail: dr_radhika_s@yahoo.co.in

Keywords

CD35, CD11b, Rapid Diagnostic tests

Abstract

Background: The Global Action Plan to combat antimicrobial resistance highlights the need for sustainable investment in diagnostic development. A major cause of irrational antibacterial prescribing is the absence of rapid tests to accurately identify pathogens. Neutrophil receptor expression has recently been proposed to differentiate bacterial from viral infections through estimation of the Clinical Infection Score (CIS). This study evaluated the feasibility and acceptability of estimating CIS operationally in clinical settings.

Methods: In Phase I, operational feasibility was assessed in healthy participants, while in Phase II, operational acceptability was assessed in patients with suspected infection who had not yet received empiric antimicrobials. Operational feasibility of CIS estimation was assessed based on successful completion of all required steps of score calculation and acceptability was assessed in terms of operational convenience from obtaining consent to calculation of CIS score. To assess objectives, blood samples were collected to measure absolute neutrophil count, erythrocyte sedimentation rate, C-reactive protein, and mean fluorescence intensity of CD35-FITC (CR1) and CD11b-PE (CR3). CIS was calculated using the algorithm proposed by Lilius and Nuutila.

Results: Phase I was done in 12 healthy participants. In phase 2, 36 patients were enrolled but only samples of 30 patients were included in analysis as 6 samples got lysed. The time from sample collection to reporting CR1 and CR3 values was 1.5–2 hours, and 1.75–2.25 hours for CIS calculation. Among healthy subjects, CIS ranged from 0–2, while in patients, 20 had CIS <2 and 10 had CIS ≥2, consistent with previous reports.

Conclusion

The findings demonstrated that measuring CR1 and CR3 expression and computing CIS is feasible operationally in healthcare settings. The CIS range in healthy controls and patients with suspected infection was consistent with findings reported in the earlier studies. However, randomized controlled studies are needed to validate the diagnostic utility of CIS and to assess its impact on reducing empirical antimicrobial therapy.

Introduction

Antimicrobial resistance (AMR) was recognized by the World Health Assembly in 2014 as a major global health threat, leading to the 2015 Global Action Plan that emphasizes sustainable investment in developing new antimicrobials, diagnostics, and vaccines [1].

The limited investment in developing new antimicrobials makes preserving the effectiveness of existing drugs increasingly urgent [1]. Although antibacterials target bacterial infections, similar symptoms across infections caused by different pathogens and the ≥48-hour turnaround of the gold-standard culture-susceptibility test often lead to empirical prescribing of antibacterials based on clinical suspicion [2].

A key driver of irrational antibacterial prescribing is the lack of rapid diagnostic tests to identify pathogens and their antimicrobial susceptibility at the point of care, limiting timely, appropriate treatment decisions [3,4].

There has been increasing focus on the use of acute-phase reactants (APRs) like erythrocyte sedimentation rate (ESR), c-reactive protein (CRP), procalcitonin (PCT) in the management of infections. However, they are known to get elevated in non-infectious, inflammatory conditions, thereby are associated with risk of false positive and false negative results for diagnosis of infections [5,6,7].

Lilius EM. and Nuutila J. proposed that neutrophil receptor expression can help differentiate bacterial from viral infections. Complement receptors CR1 (CD35) and CR3 (CD11b) are weakly expressed on resting neutrophils but are upregulated in bacterial infections; in viral infections, CR1 is downregulated while CR3 remains upregulated. They suggested that a clinical infection score (CIS), based on ESR, CRP, mean fluorescence intensity (MFI) of CR1, and the total neutrophil complement receptor (TNCR) index, can distinguish microbiologically confirmed bacterial from viral infections at a cutoff >2, with 98% sensitivity and 97% specificity [8].

Therefore, given the need for novel biomarkers to distinguish bacterial from viral infections and reduce unnecessary antibacterial use and antimicrobial resistance, this exploratory study assessed the feasibility and acceptability of estimating the Clinical Infection Score (CIS) operationally in clinical settings.

Methods

This prospective study was conducted in accordance with Good Clinical Practice guidelines after obtaining Institutional Ethics Committee approval and written informed consent from all participants.

This study was conducted in two phases: phase I assessed the operational feasibility of CIS estimation using a standardized protocol adapted from method described by Lilius et al [8], and phase II evaluated operational acceptability based on ease of use, time required to calculate the score, and explored its preliminary clinical utility in patients with suspected infection.

In phase I, for assessing operational feasibility of CIS estimation, healthy adults aged ≥ 18 years of either sex who had not taken any medication in the previous two weeks were included. Individuals with any significant medical history, hepatic or renal impairment, and pregnant or lactating women were excluded.

In phase II, to assess operational acceptability of CIS, adult immunocompetent patients aged ≥ 18 years of either sex with suspected bacterial or viral infection who had not started empirical antimicrobial therapy were included. Exclusions were suspected fungal infections, medications affecting leukocyte counts, hepatic or renal impairment, and pregnancy or lactation.

Eligible participants in both phases underwent clinical examination and in phase II, clinical details and relevant investigation reports were recorded. Blood (5 mL) was collected for ESR, CRP, and Absolute Neutrophil Count (ANC). An additional 5 mL was collected in EDTA tubes for measuring CR1 and CR3 MFI, transported at 2–4 °C within 45 minutes, and processed immediately on arrival for MFI estimation.

Procedure for estimation of MFI of CR1 and CR3

EDTA blood samples were processed to measure CR1 and CR3 expression on granulocytes by flow cytometry. From each sample, 100 µL was transferred to a flow tube and incubated with 10 µL each of CD45KR, CD35-FITC, and CD11b-PE antibodies (Immunotech SAS a Beckman Coulter company) for 15–20 min in the dark at room temperature. Erythrocytes were then lysed with 500 µL Optilyse, followed by a second dark incubation (10–15 min), centrifugation at 1500 rpm for 5 min, and washing with PBS. After a final wash, events were acquired on a Beckman Coulter Navios 3-laser, 10-color flow cytometer. Granulocytes were gated on the CD45 vs. side-scatter plot, and mean fluorescence intensity (MFI) of CD35-FITC (CR1) and CD11b-PE (CR3) was obtained from histogram analysis.

All the samples were processed in only one laboratory in which two technicians underwent training from service providers to estimate MFI of CR1 and CR3.

Estimation of CIS point

The total neutrophil complement receptor (TNCR) index was calculated as the base-10 logarithm of the product of ANC, MFI of CD35-FITC (CR1), and MFI of CD11b-PE (CR3) [8].

For each variable, values below the cutoff (CRP 77 mg/L, ESR 28 mm/h, CR1 MFI 8.7, TNCR 3.4) scored 0; values between the cutoff and a second threshold (CRP 161 mg/L, ESR 42 mm/h, CR1 MFI 13.5, TNCR 3.9) scored 1; and values above the second threshold scored 2. The CIS was calculated by

summing the scores for CRP, ESR, CR1 MFI, and TNCR index [8].

Operational feasibility of CIS estimation

Operational feasibility of CIS estimation was assessed based on successful completion of all required steps of score calculation, including blood sample collection and processing, measurement of MFI of CR1 and CR3 by flow cytometry using standardized test protocol, derivation of the TNCR index using algorithm and calculation of the final CIS using predefined cut-offs as proposed by Lilius et al [8].

Operational acceptability of CIS

Operational acceptability of CIS estimation was assessed in terms of practical implementation, including patient consent for sample collection; sample stability during transport; ease of procurement of reagents, setting up of flow cytometer for measurement of MFI of complement receptors; simplicity of CIS calculation; turnaround time from sample collection to final score estimation and cost of the test.

In addition to operational acceptability, the preliminary clinical utility was also explored by comparing CIS of cases with CIS of healthy participants.

Statistical analysis

Demographic and clinical characteristics were summarized as mean ± SD for quantitative variables. The TNCR index and CIS were calculated for all participants, and mean ± SD values were determined for healthy subjects, cases with CIS <2, and cases with CIS ≥2. Between group comparison of quantitative variables was done by One-Way ANOVA followed by pairwise comparisons.

A feasibility sample of 12 healthy subjects and 36 patients with suspected bacterial or viral infection was planned. Assuming a 20% screen failure rate, 57 subjects were required to recruit 48 participants. As this was the first study that explored the feasibility and acceptability of CIS estimation, the sample size was calculated based on feasibility.

Results

In phase I, the flow cytometer was successfully configured for estimation of MFI of CD35-FITC (CR1) and CD11b-PE (CR3) in a single sitting by the instrument service providers following procurement of the reagents. However, procurement of reagents through the institutional rate contract process required approximately two months, which delayed assay establishment. Once the assay was established, 15 healthy adults were screened; 12 males were included and 3 excluded due to recent medication use. The median age of participants was 28 year. All recruited participant samples were collected, transported at 2–4 °C to the laboratory. The time from sample collection to delivery to the laboratory ranged from 0.5 to 0.75 h. Once received at the laboratory, all the samples were processed and test execution was carried out according to the procedure briefed in methodology and MFI of CR1 and CR3 were obtained in all participants. The turnaround time from sample collection to reporting of CR1 and CR3 MFI was 1.5–2 hours. Mean CR1 and CR3 values were observed to be 5.4 and 53.7, respectively (Table 1).

The TNCR index and CIS were computed for all samples according to the algorithm proposed by Lilius et al without any missing data or assay failure. CIS calculation required approximately 1.75–2.25 hours from sample collection. In healthy subjects, CIS ranged from 0–2; 11 of 12 had CIS <2, and one had CIS =2 (Figure 1). The 11/12 obtained values were within the range previously reported in the study by Lilius et al [8]. No major logistical or technical challenges were encountered after assay establishment, indicating that implementation of CIS assay was feasible in our setting.

In phase II, 41 patients were screened and 36 enrolled. Three were excluded due to recent antimicrobial use and two declined consent. Among the 36 cases, suspected diagnoses included acute undifferentiated fever (17), upper respiratory tract infection (12), pneumonia (2), malaria (2), acute gastroenteritis (2), and dengue (1). However, in one female inpatient with clinical suspicion of pneumonia, sputum culture reported *Klebsiella pneumoniae* and one female patient with clinical suspicion of dengue was found positive for NS1 antigen. Rest all the cultures were reported sterile. Blood samples from all eligible cases were collected, transported, processed and analyzed according to the standardized protocol established in phase 1.

Six (6) samples were excluded due to lysis, resulting in successful CIS estimation in 30 of 36 cases (83.3%) with median age of 30 years (19 males and 11 females). CIS point ranged from 0–4 with 20 cases having a CIS <2 and 10 cases having CIS ≥2 (Figure 2). Mean CR1 and CR3 were 5.7 and 62.1 in cases with CIS <2, and 8.3 and 86.9 in cases with CIS ≥2 (Table 1). The CIS score of patient with *K pneumoniae* infection was 3 and for patient with dengue was 0. A dot plot of healthy subjects and cases is shown in Figure 1 and 2 respectively.

Figure 1: Showing a dot plot of CIS scores in healthy subjects.

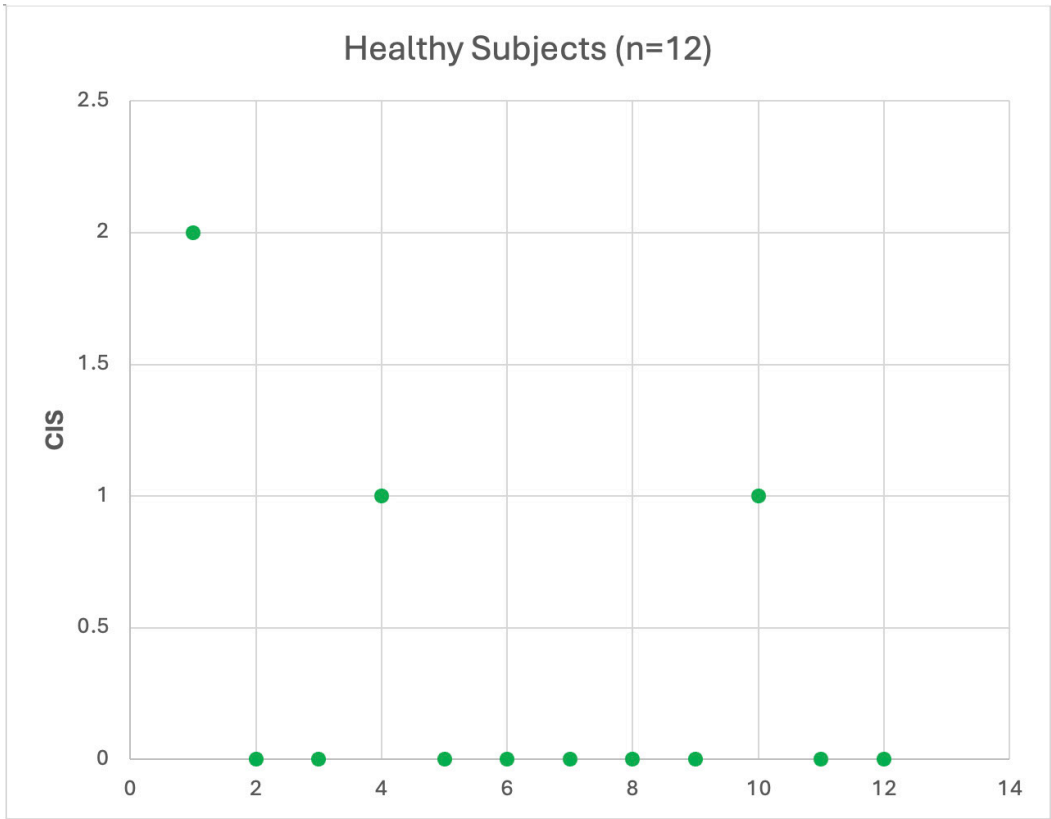
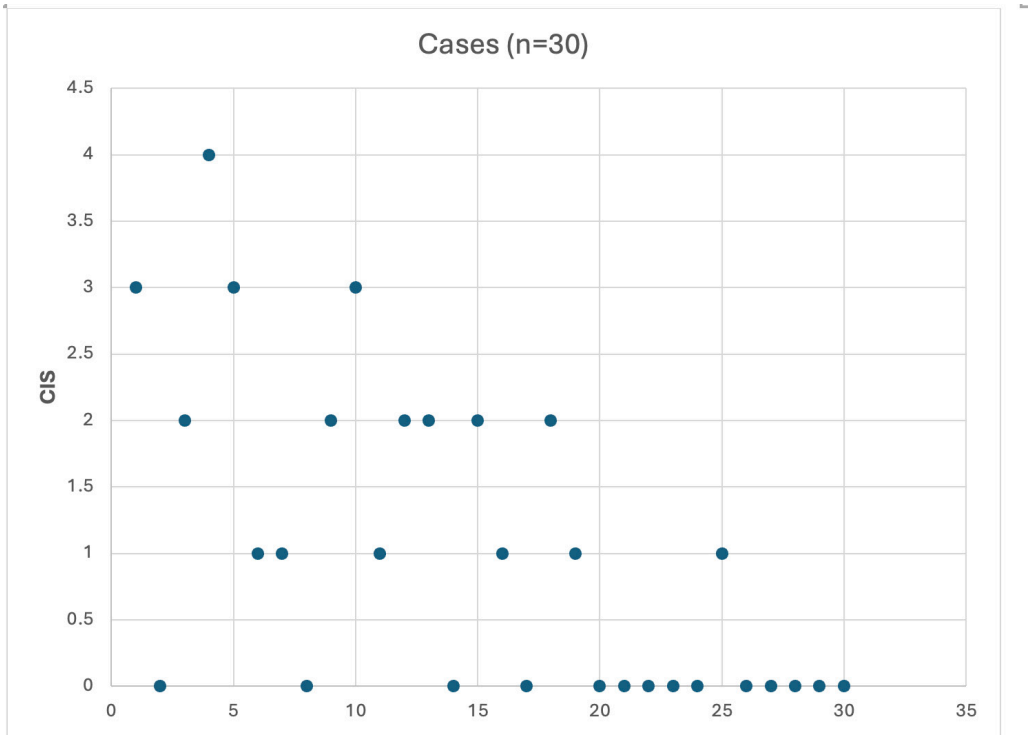


Figure 2: Showing a dot plot of CIS scores in cases.



The mean and standard deviation of ANC, ESR, CRP, MFI of CR1, MFI of CR3, TNCR index, and CIS for healthy subjects, cases with CIS <2, and cases with CIS ≥2 are presented in Table 1.

Table 1: Showing the Mean (SD) of variables in healthy subjects, cases with CIS ≥ 2 and CIS<2.

Variables	Healthy Subjects (n=12)	Cases with CIS ≥ 2 (n=10)	Cases with CIS<2 (n=20)	pvalue (One Way ANOVA)
ANC	4.6 (1.2)	4.7 (2.1)	3.6 (1.8)	0.17
ESR (mm/h)	5.7 (2.9)	32 (19.6)**	14.7 (9.5)#	<0.001
CRP (mg/L)	8 (6)	22.1 (32.1)	8.9 (8.4)	0.09
Neutrophil CR1	5.4 (1.4)	8.3 (2.6)*	5.7 (1.6)#	<0.01
Neutrophil CR3	53.7 (22.6)	86.9 (25.5)*#	62.1 (36.4)#	0.04
TNCR Index	3.1 (0.3)#	3.4 (0.5)^	2.9 (0.5)	<0.05
CIS	0.33 (0.7)#	2.5 (0.7)**	0.3 (0.5)	<0.001

*# p>0.05; ^p<0.05; *p<0.01; **p<0.001.*

When tested using One-Way ANOVA, ANC and CRP values were found to be comparable between the healthy controls, cases with CIS<2 and cases with CIS≥2 . While ESR, MFI of CR1 and CR3, TNCR index and CIS were observed to be statistically significantly different between the groups. However, the values of ESR, MFI of CR1, TNCR index and CIS were found to be significantly higher in patients with CIS≥ 2 when compared to patients with CIS<2 and healthy subjects. While MFI of CR3 was higher in patients with CIS≥ 2 than in patients with CIS<2, it was not statistically significant. In addition, it was observed that there was no significant difference in MFI of CR1, MFI of CR3, TNCR index and CIS between healthy subjects and patients with CIS<2. (Table 1)

Discussion

This exploratory study demonstrated the operational feasibility and acceptability of CIS estimation in a tertiary healthcare setting. CR1 and CR3 expression on granulocytes could be successfully measured by flow cytometry, CIS could be calculated within a clinically relevant turnaround time and in addition the findings provided preliminary evidence of clinical utility.

As the CIS is based on a blood sample, patient acceptability appeared favorable as only two patients declined consent. However, the requirement for transport at 2–4 °C and within 45–60 minutes of sample collection may limit operational flexibility and could present logistical challenges particularly when samples are analyzed at an external facility.

In this study, the MFI of CR1 was observed to be significantly higher in patients with CIS ≥ 2 than in patients with CIS< 2 and MFI of CR3 was found high in patients with CIS≥ 2 when compared to patients with CIS<2 though not statistically significant. (Table 1) These results agree with Lilius EM and Nuutila study, in which they reported that in neutrophils, CR1, and CR3 were upregulated in bacterial infections while CR1 was downregulated and CR3 was upregulated in viral infections [8]. A review by Kapasi et al. reported that among cell-surface biomarkers for differentiating bacterial from viral infections, CD64 (cutoff 1.64; up to 95.2% specificity) and CD35 (cutoff ≥1.5; up to 77% specificity) were the most reliable. The review also noted that combining markers improved diagnostic performance, including CD34/CD64; CD35+CD55 and CD35+CD32+CD88+MHC [7].

However, Nuutila reported that FcγRI (CD64) expression on neutrophils and monocytes increased significantly in both viral and bacterial febrile infections compared with healthy controls. Its sensitivity (73% and 52%) and specificity (65% and 52%) for distinguishing bacterial from viral infections were low, leading to the conclusion that CD64 is better suited as a marker of febrile infection rather than for differentiating infection aetiology [8]. This contrasts with findings reported by Kapasi et al [7].

A study from China by Lu et al. evaluated CD35 and CD64 expression on neutrophils, monocytes, and lymphocytes using flow cytometry to differentiate bacterial and viral infections in children. The NCD35/NCD64 ratio showed 90.0% sensitivity and 93.7% specificity for diagnosing viral infection at a cutoff <7.256. Individually, CD64 and CD35 had lower sensitivity and specificity than the ratio [9].

The CR1 and CR3 could be estimated in 1- 1.5 h after receiving samples and CIS can be provided to physicians within 1.75–2.25 hours of sample collection. Additionally, since samples must reach the laboratory within 1 hour, results can still be delivered within 2.25 hours even if testing is performed at an external facility.

In healthy subjects, the CIS was 0 in 9 of 12 participants, 1 in two participants, and 2 in one participant (Figure 1), consistent with values reported by Lilius E. M. et al., [8].

In the present study, the observed CIS ranged from 0 to 4, whereas Lilius EM and Nuutila reported a wider CIS range of 0 to 8 [8]. This difference may be attributed to variations in the severity of infections among the study populations. In the

current study, patients were recruited from outpatient services and general wards, with no patients from intensive care units. In contrast, in the study by Lilius EM et al., 36 out of 46 patients had severe infections such as sepsis or pyelonephritis [8].

Probably due to the inclusion of predominantly mild-to-moderate infections in the present study, the microbiological burden may have been low, resulting in a higher proportion of sterile cultures. However, it is in such patients with mild-to-moderate infections that empirical antimicrobial therapy is most commonly prescribed as culture confirmation is rarely possible, thereby further highlighting the clinical relevance and importance of such tests. Even in Lilius EM and Nuutila study, only in 84/492 (17%) of patients, infection could be confirmed microbiologically [8].

CIS ≥2 in 10 of 30 cases, suggest a likely bacterial infection (Figure 2) and CIS <2 in rest of the 20 cases suggest low likelihood of bacterial infection [8], as per the cutoff points proposed in Lilius EM and Nuutila study. However, these assumptions remain presumptive and require confirmation through appropriately designed diagnostic validation studies.

In the study of Lilius et al., 60 healthy controls and 292 hospitalized febrile patients were enrolled, of whom 46 were found to have microbiologically confirmed bacterial and 38 had microbiologically confirmed viral infection. It was reported that, in these patients at a cutoff point of >2, CIS differentiated bacterial from viral infections with 98% sensitivity and 97% specificity [8]. Whereas, procalcitonin, the most widely used marker for diagnosis of bacterial infections was reported to have a sensitivity of 76% and specificity of 69% at the commonly used cut-off of 0.5 ng/mL in a systematic review and meta-analysis assessing the diagnostic accuracy of procalcitonin for bacteraemia. The highest specificity (78%) was reported in immunocompromised/neutropenic patients, while the lowest (55%) in localized infections. The authors concluded that procalcitonin has fair diagnostic accuracy for bacteraemia in hospitalized adults with suspected infection or sepsis [10].

In a retrospective study involving 431 febrile patients with solid tumours and suspected infection, procalcitonin was estimated and culture samples were obtained before initiation of empirical antimicrobial therapy. Among 431 blood cultures, 181 were positive while the remainder were sterile. At a procalcitonin cut-off of 0.5 ng/dL, only 55% had positive blood cultures. Even using a higher cut-off of 1.52 ng/dL, procalcitonin demonstrated an AUC of 0.7 for diagnosing Gram-negative bacterial infections, with sensitivity and specificity of 72.1% and 70.1%, respectively. However, for Gram-positive/fungal infections, procalcitonin showed a statistically non-significant low AUC of 0.5, with sensitivity and specificity of 49.8% and 50.2%, respectively [11].

A cross-sectional study from Pakistan comparing the diagnostic accuracy of procalcitonin, C-reactive protein (CRP), total leukocyte count (TLC), and lactate for sepsis diagnosis included 126 adults with suspected sepsis. Using blood culture as the reference standard, ROC analysis showed that procalcitonin at a cut-off of >0.5 ng/mL had an AUC of 0.781, whereas CRP (>10 mg/L), TLC (>10 × 10^9/L), and lactate (>2.2 mmol/L) had AUCs of 0.568, 0.617, and 0.545, respectively. The study also reported a negative predictive value of 80.8% for procalcitonin at a cut-off of 0.15 ng/mL, suggesting its potential utility in ruling out sepsis when levels are <0.15 ng/mL [12].

In a similar retrospective study from China evaluating the utility of procalcitonin for diagnosing bacterial infections in patients with chronic gouty arthritis, 66 patient records were analyzed. Among these, 21 patients had microbiologically confirmed infections, while 43 had sterile cultures. Procalcitonin at a cut-off of 0.5 ng/mL showed an AUC of 0.526, with sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio of 22.2%, 61.5%, 0.577, and 1.265, respectively. The AUCs for ESR (>20 mm/h), CRP (>8 mg/L), WBC count (>10 × 10^9/L), and neutrophil ratio (>70%) were 0.53, 0.635, 0.577, and 0.712, respectively. The authors concluded that procalcitonin may not be a reliable marker for diagnosing bacterial infection in febrile patients with chronic gouty arthritis [13].

As the CIS cut-off value of >2 has been reported to demonstrate a sensitivity of 98% and specificity of 97% for differentiating bacterial from viral infections [8], along with a rapid turnaround time of approximately 2 hours, it appears to be a promising tool in the current clinical scenario, particularly when compared with other acute phase reactants for which lower sensitivity and specificity have been reported in the above-mentioned studies. However, since estimation of complement receptor expression on neutrophils is not routinely performed in clinical practice, the present study was undertaken to explore the feasibility and acceptability of CIS estimation prior to conducting a validation study.

Another advantage of the CIS is its cost-effectiveness compared with rapid pathogen identification tests such as the BioFire Film Array. A single BioFire test costs about $200 (~₹18,150), whereas the CIS costs ₹2,000–2,500 per test—an approximately 80% cost reduction—making it a more affordable option [14].

The strengths of the study are that this is the first study that explored the feasibility and acceptability of estimating CIS, a promising tool in the current clinical scenario. In this study, the range of CIS was assessed both in healthy participants and patients with suspected bacterial and viral infections.

However, the limitations of the study are that though the samples where appropriate were sent for culture susceptibility tests and viral serology/antigen tests for confirmation of aetiological agent, as most of the patients had mild-moderate infection, the tests were reported sterile/negative except in 2 cases. Therefore, in this study, inflammatory conditions/ fungal causes could not be ascertained which could have resulted in similar clinical picture in the patients.

Conclusion

To our knowledge, this is the first study that explored the feasibility and acceptability of estimating CIS. It showed that measuring CR1 and CR3 expression on granulocytes and calculating the CIS is feasible operationally in healthcare settings. The CIS range in healthy controls and patients with suspected infection was consistent with findings reported by Nuutila. However randomized controlled studies are needed for validation of CIS as rapid diagnostic test and for assessment of its impact on reducing empirical antimicrobial therapy.

Conflict of interest

All the authors declare that they have no conflict of interest. The study is done only for academic interest.

Ethics Committee Approval

The study was initiated only after obtaining Institutional Ethics Committee Approval and participants both healthy subjects and patients were recruited into the study only after obtaining written informed consent.

Declaration

The corresponding author declares that all authors on this publication have met all of the following criteria of authorship.
All the authors have
(a) significant contributed to the concept and design/ data acquisition/ analysis and interpretation of data
(b) drafted or revised the article for scientific content
(c) approved the manuscript content
(d) agreed to be accountable for all aspects of the study and manuscript preparation.
Nobody who qualifies for authorship has been omitted from the list.

Credit Author Statements

Conceptualization: Prashanth Mada, Radhika Soanker
Data curation: Prashanth Mada
Formal analysis: Radhika Soanker
Investigation: Shyamala Ravikoti, Shantveer G Uppin, Tara Roshni Paul
Methodology: Prashanth Mada, Sakthivadivel Varatharajan, Gaddala Mary Sowjanya
Supervision: Prashanth Mada, Shantveer G Uppin
Validation: all authors
Visualization: Radhika Soanker, Prashanth Mada
Writing–original draft: Radhika Soanker
Writing–review & editing: all authors.

Acknowledgements

Tha authors would like to thank Dr. Dilip Mathai, who served as Dean and Professor at The Apollo Medical College, Hyderabad and as Professor of Medicine at Christian Meidcal College, Vellore for the constant encouragement to pursue this research.

Funding

This work was supported by Intramural Research Grant from All India Institute of Medical Sciences, Bibinagar, Hyderabad Metropolitan Region. The fund was utilized for doing investigations of participants.
Grant No # AIIMS/BBN/IM-F/2022/19

Data Availability Statements

The data shall be shared upon request to corresponding author

Supplemental material

None.

Equipment used in study

Beckman Coulter Navios 3-laser, 10-color flow cytometer.

Supplier of reagents

IMMUNOTECH SAS a Beckman coulter company 13276 Marseille Cedex 9 France is the supplier for the three key reagents used in study, CD35 FITC, CD11b PE and CD45 Krome Orange.

References

1. World Health Organization. WHO report on surveillance of antibiotic consumption: 2016–2018 early implementation. Geneva: World Health Organization; 2018. Available from: https://iris.who.int/server/api/core/bitstreams/efb4b49e-ecaa-4eb4-8b1d-d225b9923696/content [Accessed on 30 Aug 2026]
2. Simon L, Gauvin F, Amre DK, Saint-Louis P, Lacroix J. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: a systematic review and meta-analysis. Clin Infect Dis. 2004;39(2):206-217. doi: 10.1086/421997.
3. Machowska A, Lundborg CS. Drivers of irrational use of antibiotics in Europe. Stockholm: Department of Public Health Sciences, Karolinska Institutet; 2018. Available from: https://haiweb.org/storage/2018/03/Report-Drivers-of-Irrational-Use-of-Antibiotics.pdf [Accessed on 07 Jan 2026].
4. Bansal A, Gupta M, Shafiq N, Saroch A, Muralidharan J, Arora N, et al. Rational use of antibiotics at primary & secondary healthcare facilities in district Panchkula, Haryana, North India: A cross-sectional study. Indian J Med Res. 2025;162(2):175-185. doi: 10.25259/IJMR_1146_2025.
5. Schuetz P, Albrich W, Mueller B. Procalcitonin for diagnosis of infection and guide to antibiotic decisions: past, present and future. BMC Med. 2011;9:107. doi: 10.1186/1741-7015-9-107.
6. Markanday A. Acute phase reactants in infections: evidence-based review and a guide for clinicians. Open Forum Infect Dis. 2015;2(3):ofv098. doi:10.1093/ofid/

ofv098.
7. Kapasi AJ, Dittrich S, González IJ, Rodwell TC. Host biomarkers for distinguishing bacterial from non-bacterial causes of acute febrile illness: a comprehensive review. PLoS One. 2016;11(8):e0160278. doi:10.1371/journal.pone.0160278.
8. Lilius EM, Nuutila J. Bacterial infections, DNA virus infections, and RNA virus infections manifest differently in neutrophil receptor expression. ScientificWorldJournal. 2012;2012:527347. doi: 10.1100/2012/527347.
9. Lu L, Jin X, Zhang Q. CD35 and CD64 of neutrophils can differentiate between bacterial and viral infections in children by simultaneous quantitative analysis. Med Sci Monit. 2019;25:7728-7734. doi:10.12659/MSM.914527.
10. Hoeboer SH, van der Geest PJ, Nieboer D, Groeneveld AB. The diagnostic accuracy of procalcitonin for bacteraemia: a systematic review and meta-analysis. Clin Microbiol Infect. 2015;21(5):474-481. doi: 10.1016/j.cmi.2014.12.026.
11. Vincenzi, B. Fioroni I, Pantano F, Angeletti S, Dicuonzo G, Zoccoli A, et al. Procalcitonin as diagnostic marker of infection in solid tumors patients with fever. Sci Rep. 2016;6:28090. doi: 10.1038/srep28090.
12. Bibi A, Basharat N, Aamir M, Haroon ZH. Procalcitonin as a biomarker of bacterial infection in critically ill patients admitted with suspected sepsis in intensive care unit of a tertiary care hospital. Pak J Med Sci. 2021;37(7):1999-2003. doi: https://doi.org/10.12669/pjms.37.7.4183
13. Zhang J, Zhao C, Wu T, Su J, Wu X, Liu J et al. Procalcitonin may not be a differential diagnostic marker for bacterial infection in febrile patients with chronic gouty arthritis. Journal of International Medical Research. 2018;46(10):4197-4206. doi:10.1177/0300060518791093
14. Fleischer E, Aronson PL. Rapid diagnostic tests for meningitis and encephalitis-BioFire. Pediatr Emerg Care. 2020;36(8):397-401. doi:10.1097/PEC.0000000000002180.

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

Evaluation of Pediatric Pleural Effusion: Insights from a Tertiary Care Center in North India

Ashna Kumar[1], Sugandhi Malgotra[2], Keshav Mehra[3], Preeti Sharma[4], Mohd. Razaq[5], Nikhil Rajvanshi[6*]

[1]*Junior Resident, Department of Pediatrics, Government Medical College, Jammu, J&K, India*
[2]*Junior Resident, Department of Radiology, Government Medical College, Jammu, J&K, India*
[3]*Junior Resident, Department of Pediatrics, Dayanand Medical College, Ludhiana, Punjab, India*
[4] *Senior Resident, Department of Pediatrics, Government Medical College, Jammu, J&K, India*
[5]*Associate Professor, Department of Pediatrics, Government Medical College, Jammu, J&K, India*
[6] *Consultant Pediatric Pulmonologist, Madhukar Rainbow Children's Hospital, New Delhi, India*

Article Info

[*]***Corresponding Author:***
Nikhil Rajvanshi
Consultant Pediatric Pulmonologist, Madhukar Rainbow Children's Hospital, New Delhi, India
E- mail: nikhilrajvanshi05@gmail.com
ORCID ID: 0000-0002-0787-8815

Keywords

Children, Empyema, Inter-costal drainage, Pleural Effusion, Tuberculosis

Abstract

Background: Pediatric pleural effusion, characterized by abnormal fluid accumulation in the pleural space, arises from various etiologies, including infections and systemic diseases. This study evaluated the clinical, radiological, and bacteriological profiles of pleural effusion and related outcomes in children in North India.

Methodology: A prospective observational study was conducted on 133 children aged 0–19 years with confirmed pleural effusion at a tertiary care centre. Comprehensive clinical history, socioeconomic data, physical examinations, imaging, and pleural fluid (PF) analysis were performed. Microbiological testing, including CBNAAT for Mycobacterium tuberculosis, was utilized when tuberculosis was suspected. Follow-ups were conducted until 6 weeks post-discharge.

Results: The mean age was 8.67 ± 4.93 years, with a male predominance (54.1%). Common symptoms included fever (85.7%) and cough (67.7%), and right-sided effusion (49.6%) was most frequent. Diagnoses included tuberculosis (21.1%), bacterial empyema (20.3%), and sepsis (16.5%). PF cultures detected Staphylococcus aureus in 3.8% and Streptococcus pneumoniae in 2.3%. CBNAAT was positive in only 7.1% of PF samples but in 57.1% of respiratory specimens (sputum or gastric aspirate). Most cases (62.4%) were managed conservatively, 20.3% required intercostal tube drainage, and 10.5% underwent surgery. The median ICT duration was 239 hours, and median hospital stay was 342 hours. Mortality was 6.8%, mainly from sepsis.

Conclusion: This study underscores diverse etiologies of pediatric pleural effusion, emphasizing tuberculosis and empyema. Timely diagnosis and management, including surgical intervention when needed, are essential to enhance outcomes.

Introduction

Pleural effusion results from an abnormal accumulation of fluid in the pleural space, often due to an imbalance between fluid production and drainage [1]. In children, the leading causes include infections, congestive heart failure, and nephrotic syndrome. Pneumonia often serves as a common precursor, accounting for up to 70% of cases [1]. Some cases of pneumonia can progress to empyema, which occurs in approximately 0.6-2% of children with bacterial pneumonia [2]. The most frequently isolated organisms in such cases include Staphylococcus aureus, followed by Streptococcus pneumoniae and Klebsiella [3,4]. The prognosis of pleural effusion in children largely depends on the underlying disorder and the treatment administered [5]. In India, particularly within the pediatric population, factors such as delayed medical consultation and the indiscriminate use of antibiotics contribute to poor diagnostic yields, leading to increased morbidity from complications like empyema [6].

Tuberculous effusions are generally observed in adolescents, whereas empyema is more prevalent among children aged 1 to 5 years, with a higher incidence in males and those from lower socio-economic backgrounds [5,7,8]. Predisposing factors such as cerebral palsy, congenital heart disease, and immunosuppression are commonly associated with pleural effusion; although it can also occur without any underlying conditions [9]. Differentiating between exudative and transudative effusions using Light's criteria can aid in identifying the underlying pathology [10,11]. Imaging modalities, including chest radiographs, ultrasound, and, in specific cases, contrast-enhanced computed tomography (CECT), play a crucial role in diagnosis and management [12,13,14].

Given the diverse etiologies and outcomes associated with pleural effusion, understanding the clinical, radiological, and bacteriological profiles is essential for optimizing management strategies. The high incidence of complications such as empyema, combined with regional variations in presentation and outcomes, highlights the need for localized data to inform treatment approaches. Therefore, this study was designed to elucidate the clinical, radiological, and bacteriological profiles of pleural effusion in hospitalized children in North India. Additionally, the study aimed to evaluate the outcomes related to the type and etiology of pleural effusion in the pediatric age group.

Material and Methods

This prospective observational study was conducted at a tertiary care center in North India after obtaining approval from the Institutional Ethics Committee(IEC/GMC/2019/787). The study included all the children aged 0 to 19 years admitted with clinical and radiological evidence of pleural effusion during the study period. The upper age limit of 19 years was adopted because, at our institution, patients up to 19 years of age are routinely managed under the Department of Pediatrics, in accordance with institutional admission policy Exclusion criteria included those who either did not provide informed consent or were lost to follow-up during the course of the study.

Data Collection

For each enrolled patient, a detailed clinical history was taken, which included information on symptoms, their duration, and any relevant medical history, such as prior infections, tuberculosis contact, or comorbidities like congenital heart disease, immunosuppressive conditions or any underlying chronic disease. Socio-economic status as per modified Kuppuswamy scale, was recorded for all participants to examine its association with disease outcomes. A thorough physical examination was performed for all patients.

Diagnostic Investigations

Both chest X-rays and ultrasound imaging were performed for all patients to confirm the diagnosis of pleural effusion and assess fluid characteristics such as size, location, and any associated conditions like lung consolidation or collapse. Ultrasound also provided guidance for pleural fluid (PF) aspiration when required. In cases of minimal effusion without respiratory compromise, PF aspiration was avoided to minimize unnecessary invasive procedures.

Computed tomography (CT) scans were reserved for cases where detailed anatomical evaluation was necessary, especially when surgical intervention was planned or if complications like loculated effusions, lung abscesses, or other associated anomalies were suspected, based on the clinician's discretion. PF aspiration was performed under aseptic conditions, and the aspirated fluid underwent biochemical (protein, glucose, lactate dehydrogenase), cytological (cell count and differential), and bacteriological testing. Light's criteria were used to classify the pleural effusion as either exudative or transudative. PF pH was measured using a blood gas analyzer. Samples were also sent for Gram stain, bacterial cultures, and polymerase chain reaction (PCR) for detecting Mycobacterium tuberculosis (M.tb) through CBNAAT (Cartridge Based Nucleic Acid Amplification Test) whenever indicated. Additional diagnostic workup was done depending on suspected etiology.

Pleural fluid samples were collected under aseptic precautions and transported promptly to the central clinical laboratory for analysis. Biochemical analysis included estimation of total protein, glucose, and lactate dehydrogenase (LDH), which were performed using standard automated spectrophotometric methods in accordance with the manufacturer's instructions and the institutional laboratory's internal quality control procedures. Simultaneous blood samples were obtained, wherever indicated, to facilitate the application of Light's criteria.. Pleural fluid samples for pH estimation were collected in a heparinized blood gas syringe, avoiding air bubbles, and analyzed immediately after collection to reduce pre-analytical variation. Samples intended for microbiological evaluation were processed without delay and subjected to Gram staining, bacterial culture, and

CBNAAT (Cartridge Based Nucleic Acid Amplification Test) for Mycobacterium tuberculosis (M. tb.), whenever clinically indicated, following standard laboratory protocols.
The final diagnosis was established based on the clinical presentation, radiological findings, pleural fluid analysis, microbiological investigations, and additional diagnostic tests, as appropriate. Children with pleural effusion secondary to pneumonia were classified as having either simple parapneumonic effusion or empyema. This distinction was based on thoracic ultrasonography and pleural fluid analysis. Empyema/ Complicated parapneumonic effusion was diagnosed by the presence of grossly purulent pleural fluid, positive Gram stain or bacterial culture of pleural fluid, or ultrasonographic features suggestive of a complicated pleural collection, including echogenic debris, septations, loculations, or pleural thickening (pleural peel). Pleural effusions associated with pneumonia that did not demonstrate these features were classified as simple parapneumonic effusions.

Follow-Up and Outcome Assessment

After the initial management during the hospital stay, all the patients were followed up on OPD basis four weeks and six weeks after discharge. The final outcome was recorded at six weeks with repeat chest x-ray and ultrasound thorax to assess any residual effusion or re-accumulation of PF. Any patient requiring surgical intervention was referred to surgical department and the final outcome was then recorded at 6 weeks post-discharge.

Statistical Analysis

Data were systematically entered into Microsoft Excel for organization and later analyzed using SPSS version 20.
The baseline characteristics of the study population were summarized using descriptive statistics, with mean and standard deviation or median (IQR) for continuous variables and frequency and percentage for categorical variables. The distribution of continuous variables was assessed using the Shapiro–Wilk test. Continuous variables with a non-normal distribution were compared across etiological groups using the Kruskal–Wallis test, while categorical variables were analyzed using the Chi-square test or Fisher's exact test, as appropriate Statistical significance was set at a p-value of <0.05 for all tests.

Results

A total of 133 children with pleural effusion were included in the study. The demographic and clinical characteristics are summarized in Table 1. The mean age at presentation was 8.67 ± 4.93 years with 72 (54.1%) males. Common symptoms included fever (85.71%), cough (67.66%), and shortness of breath (53.38%). The predominant clinical signs were decreased breath sounds (90.22%) and dull percussion note (87.96%). 78.2% patients were found to be anaemic.

Table 1: Baseline Characteristics of the study Population (N=133).

Characteristic	Value {mean ± SD / n(%)}
Age (years)	8.67 ± 4.93 years
Male	72 (54.1%)
Clinical features	
Fever	114 (85.71)
Cough	90 (67.66)
Chest pain	34 (25.56)
Shortness of breath	71 (53.38)
Generalized swelling	22 (16.54)
Ascites	18 (13.53)
Pedal edema	9 (6.76)
Localized swelling	5 (3.75)
Jaundice	8 (6.01)
Vomiting	24 (18.04)
Urinary complaints	11 (8.27)

Characteristic	Value {mean ± SD / n(%)}
Clinical Signs	
Tachypnea	91 (68.42)
Tachycardia	71 (53.38)
Anemia	105 (78.9)
Lymphadenopathy	34 (25.56)
Dull percussion	117 (87.96)
Bronchial breathing	15 (11.27)
Decreased breath sounds	120 (90.22)
Immunization Status	
Completely immunized	90 (67.7)
Partially immunized	14 (10.5)
Unimmunized	29 (21.8)
Socioeconomic Status (Modified Kuppuswamy Scale)	
Lower class (<10)	62 (46.61)
Middle class (1125)	46 (34.58)
Upper class (2629)	25 (18.79)

Out of the total patients enrolled in the study, 105 (78.9%) were anaemic as per the World Health Organization (WHO) criteria, with a mean haemoglobin level of 8.96 ± 1.95 gm/dL. Among these anaemic children, 17 had severe anaemia, 86 had moderate anaemia, and 2 had mild anaemia. The mean total leucocyte count (TLC) was 11,962.33 ± 6,949.28 cells/mm³. Based on age-appropriate laboratory reference ranges, 5 (3.8%) children had a decreased TLC, 75 (56.4%) had a TLC within the normal range, and 53 (39.8%) had an elevated TLC.
Chest radiographs and ultrasound revealed that right-sided pleural effusion was the most common, occurring in 66 (49.6%) children, followed by bilateral effusion in 37 (27.8%) and left-sided effusion in 30 (22.6%). Underlying lung parenchymal abnormalities were found in 69 (51.87%) children, including consolidation in 65 (48.8%) and necrotizing pneumonia in 4 (3%). Thick septation with loculation was observed in 14 (10.5%) children. CT thorax was performed in 17 (12.8%) children, providing detailed anatomical information that facilitated decortication planning in 14 (10.5%) cases and guided conservative management in 3 (2.3%) cases with necrotizing pneumonia. Additional findings on chest x-ray included mediastinal lymphadenopathy in 27 (20.3%) children, with 25 cases diagnosed as tuberculosis and 2 as haematological malignancy.
Diagnostic aspiration of PF was performed in 101(75.9%) cases, as 32 (24.06%) children had minimal effusion not amenable to tapping even with ultrasound guidance. Analysis according to Light's criteria classified 28.7% (n=29) of the effusions as transudative and 71.3% (n=72) as exudative. PF analysis showed low glucose (<60 mg/dl) in 43.5% (n=44/101) of samples, and elevated PF to serum protein ratio (>0.5), PF to serum LDH ratio (>0.6), and PF LDH (>2/3rd serum LDH) were observed in 54.4% (n=55/101), 66.3% (n=67/101), and 60.3% (n=61/101) of the samples, respectively.
Microbiological cultures of PF identified Staphylococcus aureus in 5 (3.8%) cases and Streptococcus pneumoniae in 3 (2.3%) cases. PF CBNAAT detected M.tb in only 2 of the 28 patients (7.1%) diagnosed with tuberculosis.

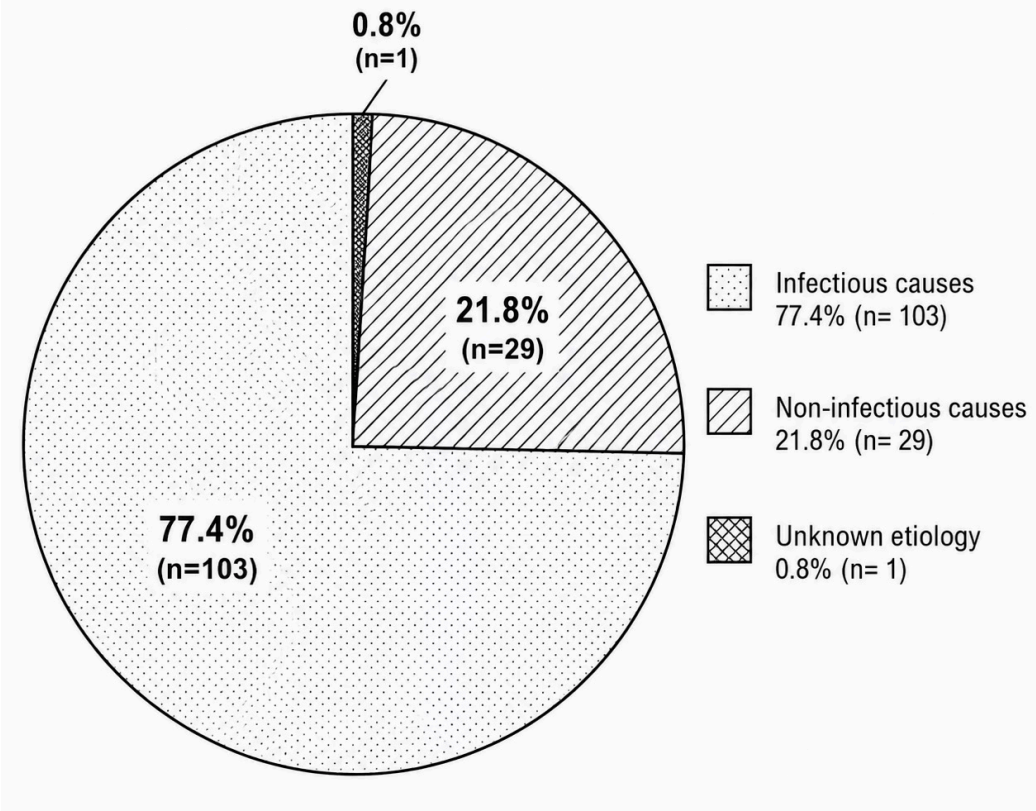
Final Diagnosis

Table 2 and Figure 1 summarize the distribution of the final etiological diagnoses based on detailed clinical evaluation.

Table 2: Final Diagnosis of Pleural Effusion Cases (N=133).

Etiological category	Diagnosis	n (%)
Infectious causes (n=103, 77.4%)		
	Tuberculosis	28 (21.1)
	Empyema (complicated parapneumonic effusion)	27 (20.3)
	Simple parapneumonic effusion	20 (15.0)
	Sepsisassociated pleural effusion	22 (16.5)
	Dengue fever	6 (4.5)
Noninfectious causes (n=29, 21.8%)		
	Nephrotic syndrome	12 (9.0)
	Acute hepatitis	11 (8.3)
	Malignancy	4 (3.0)
	Acute pancreatitis	2 (1.5)
Undetermined	Unknown etiology	1 (0.8)

Figure 1: Distribution of the etiological categories of pleural effusion in the study population (N = 133).



One child died within 4 hours of presentation in the emergency room, and a complete evaluation to determine the underlying cause could not be performed. Out of the 28 patients with tuberculosis, 16 (57.1%) had M.tb detected on CBNAAT in sputum or gastric aspirate. Among the two children with M.tb detected in PF, one had M.tb detected in both gastric aspirate and PF. Additionally, one child had M.tb detected in FNAC of cervical lymph node. The remaining 10 (35.7%) patients were diagnosed based on clinico-radiological findings.

Out of the 133 children included in the study, 83 (62.4%) were managed conservatively, while 27 (20.3%) children with empyema required intercostal tube (ICT) drainage, with fibrinolytics administered to 13 (48.1%) out of 27 children with empyema. Decortication was needed for rest of 14 (51.9%) children. The median duration of ICT drainage was 239 (168-300) hours, and the median length of hospital stay was 342 (236-508) hours. Nine (6.8%) children died due to septicaemia and multi-organ dysfunction syndrome.

Of the 124 children who survived, 9 (7.3%) were lost to follow-up. The remaining children were clinically well with no new complaints. Among those followed up, children with dengue fever, septicaemia, acute hepatitis, nephrotic syndrome and acute pancreatitis had normal chest X-rays and clear costophrenic angles on ultrasound at the 6-week follow-up.

Of the 28 children with tuberculosis, 3 (10.7%) had mild blunting of the costophrenic angle on ultrasound, and 4 (14.3%) had persistent mediastinal widening on chest X-rays. Among the 11 children with empyema who underwent decortication and completed follow-up none showed residual effusion or recurrence at 6 weeks post-discharge. Of the 27 children managed with intercostal drainage, 2 died during their hospital stay, and 3 of the remaining 25 (12%) had minimal pleural fluid at 6 weeks but were asymptomatic. Additionally, 4 out of 20 children (20%) with simple parapneumonic effusion had some pulmonary infiltrates on chest X-ray but no evidence of effusion on either chest X-ray or ultrasound.

To evaluate the relationship between clinical and pleural fluid (PF) parameters and the etiology of pleural effusion, statistical analyses were conducted to compare important variables across the different etiological groups. One child, for whom an underlying diagnosis could not be established, was excluded from this analysis. We focused on parameters such as age, hemoglobin (Hb) levels, total leukocyte count (TLC), and specific PF characteristics, given their relevance in differentiating the causes of effusion. These analyses aimed to identify significant patterns that could better inform diagnostic and treatment strategies. The findings, summarized in Table 3, highlight key associations that reflect varying disease characteristics and their clinical impacts.

Table 3: Association of various clinical or laboratory parameters with the etiology of Pleural Effusion (N=132).

Variable	Tuberculosis (n=28)	Empyema (n=27)	Simple parapneumonic effusion (n=20)	Sepsis (n=22)	Dengue (n=6)	Nephrotic syndrome (n=12)	Acute hepatitis (n=11)	Malignancy (n=4)	Acute pancreatitis (n=2)	Statistical test	p-value*
Age (years), mean ± SD	11.64 ± 4.29	7.08 ± 4.58	6.6 ± 2.0	9.6 ± 2.8	3.6 ± 4.0	12.5 ± 1.5	5.8 ± 3.2	11.7 ± 2.0	10.2 ± 1.7	One-way ANOVA	0.011
Male sex, n (%)	16 (57.1%)	12 (44.4%)	14 (70.0%)	9 (40.9%)	2 (33.3%)	8 (66.7%)	6 (54.5%)	3 (75.0%)	1 (50.0%)	Chi-square†	0.115
Duration of fever (days), median (IQR)	6 (5-8)	9 (7-12)	2 (1-4)	9 (7-13)	4 (3-7)	5 (3-7)	6 (4-9)	3 (1-5)	5 (4-7)	Kruskal-Wallis	0.018
Duration of cough (days), median (IQR)	4 (3-7)	7 (6-9)	7 (6-11)	9 (8-11)	8 (7-9)	3 (2-5)	4 (2-5)	7 (5-11)	9 (8-13)	Kruskal-Wallis	0.324
Right-sided pleural effusion, n (%)	15 (53.5%)	16 (59.2%)	11 (55.0%)	12 (54.5%)	4 (66.6%)	4 (33.3%)	3 (27.3%)	1 (25.0%)	0 (0.0%)	Chi-square†	0.005
BMI, kg/m²	16.3 ± 1.3	17.1 ± 2.3	19.2 ± 3.4	17.1 ± 3.6	19.6 ± 3.2	20.6 ± 2.6	18.1 ± 1.7	17.0 ± 1.7	18.9 ± 2.1	Kruskal-Wallis	0.491
Hemoglobin, gm/dL	11.4 ± 1.0	10.7 ± 1.9	9.6 ± 1.4	10.3 ± 1.5	10.8 ± 1.8	12.1 ± 1.8	10.7 ± 1.3	8.5 ± 2.2	11.7 ± 2.0	Kruskal-Wallis	0.002
TLC, cells/cumm	10079 (8661-11053)	18249 (10721-19537)	16546 (8126-17629)	16949 (16316-17212)	15775 (15496-16159)	10924 (9614-13282)	10106 (9915-12005)	12492 (11937-12515)	12600 (10540-14856)	Kruskal-Wallis	<0.001
PF TLC, cells/cumm	2723 (2186-3881)	4248 (3268-4348)	3480 (1864-3804)	4556 (3428-4667)	1618 (1261-2330)	2967 (2926-3187)	1515 (1392-1751)	1901 (1656-2823)	–	Kruskal-Wallis	<0.001
PF Neutrophil Count, cells/cumm	931 (923-1838)	2464 (1731-2657)	2219 (1861-2464)	1978 (1599-2261)	888 (860-894)	913 (815-1936)	1110 (622-1808)	197 (150-1148)	–	Kruskal-Wallis	<0.001
PF Lymphocyte Count, cells/cumm	1008 (967-1058)	636 (351-787)	623 (536-671)	799 (591-926)	453 (270-671)	802 (152-928)	763 (736-1864)	1371 (1062-1880)	–	Kruskal-Wallis	<0.001
PF Protein, gm/dL	4.6 ± 1.0	4.7 ± 1.2	2.5 ± 1.0	2.4 ± 1.2	3.1 ± 0.9	1.2 ± 1.4	4.2 ± 0.9	3.1 ± 1.2	–	Kruskal-Wallis	<0.001
PF Glucose, mg/dL	80 ± 20	50 ± 16	68 ± 10	74 ± 15	85 ± 8	72 ± 20	70 ± 11	47 ± 17	–	Kruskal-Wallis	0.007
PF LDH, U/L, median (IQR)	787 (526-957)	1090 (834-3335)	423 (112-853)	186 (123-288)	484 (303-664)	100 (98-102)	111 (102-123)	635 (377-892)	–	Kruskal-Wallis	<0.001
Duration of hospital stay (days), median (IQR)	12 (10-20)	18 (14-28)	06 (03-09)	15 (14-21)	6 (5-21)	08 (6-17)	5 (4-8)	27 (23-30)	12 (10-17)	Kruskal-Wallis	<0.001

Significant associations were observed between age, hemoglobin (Hb) levels, pleural fluid (PF) analysis, and the etiology of pleural effusion. Tuberculosis (TB) was more common in adolescents, while empyema affected younger children, and malignancies were prevalent in older children. The mean ages of presentation for empyema and tuberculosis were 7.08±4.58 years and 11.64±4.29 years respectively. While we observed a trend of lower BMI in conditions like empyema, TB, and sepsis, and higher BMI in nephrotic syndrome, dengue, and parapneumonic effusions, this difference did not reach statistical significance.

PF total leukocyte count (TLC), neutrophil, and lymphocyte profiles were key in distinguishing etiologies. TLC and neutrophil counts were highest in empyema and simple parapneumonic effusions, indicative of infections, while lymphocyte counts were elevated in malignancies and TB. PF protein was highest in TB and empyema, while glucose was lowest in malignancies and empyema. PF LDH levels also varied significantly across the different etiological groups, with higher values observed in infective pleural effusions.

The duration of hospital stay varied significantly with etiology, with malignancies and empyema leading to longer stays, while conditions like dengue and acute hepatitis resulted in shorter hospitalizations. These findings highlight the influence of underlying etiology and nutritional status on disease severity and clinical outcomes.

Discussion

Pleural effusion in children is a complex condition with a wide spectrum of causes, varying based on age, geography, and clinical presentation. Our study, which included 133 children, aimed to describe the clinical profile and outcomes of paediatric pleural effusion in a tertiary care setting.

In our study, tuberculosis (21.1%) and empyema (20.3%) were the most common etiologies of pleural effusion. This finding aligns with Eastham KM et al. and Efrati O et al, who reported that infectious causes dominate pediatric pleural effusion cases, particularly in low- and middle-income countries [1,15]. However, our findings differ slightly from those of Das I et al., who reported empyema as the most frequent cause in their cohort from West Bengal [3]. Similarly, in the study by Saliya MP et al. in Gujarat, parapneumonic effusion was the most common etiology, followed by TB and empyema [16]. The predominance of tuberculosis and empyema in our cohort reflects the unique epidemiological profile of pleural effusion in low- and middle-income countries. Unlike studies from high-income settings, tuberculosis continues to represent a major cause of pediatric pleural effusion in India. This finding highlights the importance of maintaining a high index of suspicion for tuberculosis, particularly among older children and adolescents presenting with pleural effusion.

Our study identified a male predominance in pleural effusion cases, with males constituting 54.1% and a male-to-female ratio of 1.18:1, similar to the observations by Eastham KM et al. and Saliya MP et al., who also noted a higher prevalence in males [15,16]. This gender predisposition may indicate a biological susceptibility in boys or differences in health-seeking behaviour, though further research is needed to explore this factor.

The age distribution in our study showed the highest incidence in children aged 6–12 years (36.8%), with a mean age of 8.6 years. These results are in line with Saliya MP et al. , who also reported a predominance of pleural effusion in older children [16]. However, Das I et al. found that the majority of cases occurred in children aged 1–5 years [3,15]. This variation may be explained by differences in study populations and local healthcare access, particularly for younger children.

The clinical presentation in our cohort was dominated by fever (85.7%), cough (67.6%), and shortness of breath (53.3%). Lingayat AM et al. reported similar symptoms in their study [8]. Das I et al and Satpathy et al. observed a highest prevalence of shortness of breath and fast breathing as a presenting complaint [3,17]. Differences in symptom frequency may be attributed to the underlying etiology and the timing of presentation to healthcare facilities. In our study, shortness of breath was more prevalent in children with empyema, which often presents with rapid respiratory compromise.

Tachypnea, dullness on percussion, and diminished breath sounds were the most prevalent clinical signs in our study. Efrati O et al. emphasized the challenge of identifying pleural effusion-specific signs in younger children due to their smaller chest surfaces [1]. Additionally, Merino JM et al. noted the overlapping physical findings in tuberculous and bacterial effusions [18]. Our findings underscore the importance of using adjunct imaging studies, such as ultrasound, to assist in the early diagnosis and management of pleural effusion, especially in younger children where physical examination alone may be inadequate.

Although not statistically significant (p=0.491), patients with empyema, TB, and sepsis tended to have lower BMI compared to those with nephrotic syndrome, dengue, and parapneumonic effusions. Nutritional status thus plays a critical role in the pathophysiology of pleural effusion, and addressing malnutrition may prevent complications in at-risk populations.

The yield of pleural fluid cultures in our study was low [n=8/101 (7.9%)], including cases of empyema, septicemia, and parapneumonic effusion. This result is comparable to Saliya MP et al.'s finding of 11.67%, though significantly lower than the 40% isolation reported by Narayanaapa et al. [16,19] The low culture yield may be attributed to prior antibiotic use, delayed presentation to tertiary care, and the persistence of inflammation even after pathogen clearance. Lingayat AM et al. similarly noted that prior antibiotic administration reduced culture positivity rates [8].The widespread administration of empirical antibiotics before referral, observed in all children in our study, is likely to have reduced the microbiological yield of pleural fluid cultures. This finding emphasizes the need for judicious antibiotic stewardship and early referral to tertiary

care centres for appropriate diagnostic evaluation.
Notably, all children in our study, irrespective of their underlying etiology, had received antibiotics prior to their hospital admission, reflecting a concerning trend of antibiotic misuse in India. Staphylococcus aureus and Streptococcus pneumoniae were the most commonly identified pathogens in our study, consistent with findings by Lingayat AM et al. and Monnier LA et al., underscoring the importance of empiric antibiotic therapy to cover these organisms when managing pleural effusion [4,8]. However, both blood and pleural fluid cultures for Mycobacterium tuberculosis were negative in all patients, a finding similar to those reported by Saliya MP et al. [16]
In our study, pleural fluid analysis highlighted total leukocyte count (TLC), neutrophil count, and lymphocyte count as key indicators for distinguishing the etiology of pleural effusion. We found that TLC and neutrophil counts were highest in cases of empyema and parapneumonic effusions, suggesting a bacterial etiology. In contrast, elevated lymphocyte counts were more commonly associated with malignancies and tuberculosis (TB). Additionally, pleural fluid protein levels were elevated in both TB and empyema, while glucose levels were lowest in cases of malignancy and empyema. These findings are consistent with those of Saliya MP et al., who also demonstrated a significant relationship between pleural fluid glucose levels and the underlying cause. Our study also demonstrated a significant association between pleural fluid LDH levels and the underlying etiology of pleural effusion. Higher LDH levels were observed predominantly in infective pleural effusions, reflecting greater pleural inflammation and cellular injury.[16]. Similar findings were observed by Das I et al., who noted correlations between neutrophil and lymphocyte counts with the underlying cause, but their study did not reveal a strong association between protein, LDH, or glucose levels with specific etiologies [3]. These findings suggest that pleural fluid biochemical and cellular parameters remain valuable adjuncts in differentiating the underlying etiology of pediatric pleural effusion, particularly in resource-limited settings where advanced molecular diagnostics may not be readily available.
The overall case fatality rate in our study was 6.8%, which is higher compared to the 1.5% reported by Lingayat AM et al. (n=1/67) and the 3.8% reported by Satpathy et al. (n=2/53) in 2015 [8,17]. However, their studies focused exclusively on children with empyema, whereas our study included all cases of pleural effusion. The majority of deaths were observed in children with septicemia and multi-system derangements, highlighting the increased severity of pleural effusion when associated with systemic infections or malignancies.
Our study's strength lies in its comprehensive prospective analysis of pediatric pleural effusion using clinical, radiological, and microbiological data, offering new insights into the etiology of the condition in Northern India. However, limitations include the single-centre design, which may not represent the broader population, and low pleural fluid culture yields due to prior antibiotic use, highlighting the need for advanced diagnostic tools. Additional limitations include the relatively small number of patients in certain etiological categories, which may have limited the statistical power for subgroup analyses. Furthermore, advanced pleural fluid biomarkers such as adenosine deaminase (ADA), interferon-γ, and molecular pathogen detection assays were not uniformly available, which may have influenced the diagnostic yield in selected cases.

Conclusion

Our study underscores the diverse etiologies of pediatric pleural effusion, with tuberculosis and empyema being the most common causes in this region. Significant associations of age and pleural fluid laboratory parameters with the underlying etiology provide valuable diagnostic insights. Despite challenges such as late referrals and low culture yields, early and appropriate management, including ICT drainage and decortication, when necessary, generally results in favourable outcomes. Notably, our study highlights the presence of various non-infectious causes of pleural effusion and reflects a concerning trend of antibiotic misuse in India, as all children had received antibiotics before hospital admission. Further research with larger cohorts and extended follow-up is needed to better understand the impact of nutritional and other risk factors on the development and outcomes of pleural effusion in children.

Institutional Ethics Committee Reference Number
IEC/GMC/2019/787.

Source of funding
Nil.

Conflict of Interest
Nil.

Acknowledgment
None.

CRediT Author Statement
The manuscript titled "Evaluation of Pediatric Pleural Effusion: Insights from a Tertiary Care Center in North India" represents a multidisciplinary effort, with each contributor fulfilling the authorship criteria established by the journal.
Dr. Ashna Kumar: Contributed to the conception of the study, data acquisition, and drafting the manuscript.
Dr. Sugandhi Malgotra: Provided critical radiological expertise for data interpretation and contributed to revising the manuscript for intellectual content.
Dr. Preeti Sharma and Dr. Keshav Mehra: Played vital roles in data acquisition, patient management, and revising the manuscript for important intellectual content.
Dr. Mohd. Razaq: Offered senior guidance in study design,

data analysis, and provided critical revisions to ensure the manuscript's quality.
Dr. Nikhil Rajvanshi: As the corresponding author and a specialist in pediatric pulmonology, was involved in interpretation of data, critical revisions, and final approval of the manuscript.
Each contributor has substantially contributed to the conception and design, drafting or revising the manuscript, and final approval of the version to be published.
Additionally, Dr. Nikhil Rajvanshi is designated as the guarantor of the integrity of the work as a whole, from inception to the published article.

References

1. Efrati O, Barak A. Pleural effusions in the pediatric population. Pediatr Rev 2002;23(12):417-426.
2. Soto-Martinez M, Massie J. Chylothorax: Diagnosis and management in children. Paediatr Respir Rev 2009;10(4):199-207.
3. Das I, Misra S, Jana PK, Sinha MK. A Study of etiology clinical profile and outcome of pleural effusion in children in a tertiary care hospital in Eastern India. Paripex-Indian Journal of Research 2017;6(4):62-65.
4. Le Monnier A, Carbonnelle E, Zahar JR, Le Bourgeois M, Abachin E, Quesne G, et al. Microbiological diagnosis of empyema in children: comparative evaluations by culture, polymerase chain reaction, and pneumococcal antigen detection in pleural fluids. Clin Infect Dis 2006;42(8):1135-1140.
5. Afsharpaiman S, Izadi M, Ajudani R, Khosravi MH. Pleural effusion in children: A review article and literature review. Int J Med Rev 2016;3(1):365-370.
6. Ball CG, Lord J, Laupland KB, Gmora S, Mulloy RH, Ng AK, et al. Chest tube complications: How well are we training our residents? Can J Surg 2007;50(6):450-458.
7. Munoz-Almagro C, Jordan I, Gene A, Latorre C, Garcia-Garcia JJ, Pallares R. Emergence of invasive pneumococcal disease caused by nonvaccine serotypes in the era of 7-valent conjugate vaccine. Clin Infect Dis 2008;46(2):174-182.
8. Lingayat AM, Wankhede PR. Study of clinical profile etiological agents and outcome in pediatric patients of empyema. Indian J Basic Appl Res 2015;4(2):502-509.
9. Singh M, Singh SK, Chowdhary SK. Management of Empyema Thoracis in Children. Indian Pediatrics 2002;39:145-157.
10. Paramothayan NS, Barron J. New criteria for the differentiation between transudates and exudates. J Clin Pathol 2002;55:69-71.
11. Light RW, Macgregor MI, Luchsinger PC, Ball WC. Pleural effusions: the diagnostic separation of transudates and exudates. Ann Intern Med 1972;77(4):507-513.
12. Avansino JR, Goldman B, Sawin RS, Flum JR. Primary operative versus nonoperative therapy for pediatric empyema: A meta-analysis. Pediatrics 2005;115(6):1652-1659.
13. Qureshi NR, Gleeson FV. Imaging of pleural disease. Clin Chest Med 2006;27(2):193-213.
14. Davies CWH, Gleeson FV, Davies RJO. BTS guidelines for management of pleural infection. Thorax 2003;59(suppl II):ii18-28.
15. Eastham KM, Freeman R, Kearns AM, Eltringham G, Clark J, Leeming J, et al. Clinical features, etiology, and outcome of empyema in children in the North East of England. Thorax 2004;59(6):522-25.
16. Saliya MP, Joshi GS. Profile of children with pleural effusion in an urban tertiary care hospital. Int J Contemp Pediatr 2017;4:1857-1860.
17. Satpathy SK, Behera CK, Nanda P. Outcome of parapneumonic empyema. Indian J Pediatr 2005;72(3):197-199.
18. Merino JM, Carpintero I, Alvarez T, Rodrigo J, Sanchez J, Coello JM. Tuberculous pleural effusion in children. Chest 1999;115(1):26-30.
19. Narayanappa D, Rashmi N, Prasad NA, Kumar A. Clinico-bacteriological profile and outcome of empyema. Indian Pediatr 2013;50(8):783-785.

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

Evaluation of Anti-HCV, HBsAg and HAV IgM on Three Different Chemiluminescence Immunoassay Systems for Hepatitis Screening

Diana G. Coste[1*], Stefan Dascalu[1,2], Anthony J. Nagle[1]

[1]*Dig Worldwide Ltd, Innovation House, Sandwich, United Kingdom*
[2]*Department of Biology, University of Oxford, Oxford, United Kingdom*

Article Info

[*]***Corresponding Author:***
Diana G. Coste
Dig Worldwide Ltd, Innovation House, Sandwich, United Kingdom
E-mail: dgc@dig-worldwide.com

Keywords

Analytical sensitivity, Anti-HCV; Chemiluminescent immunoassay, CLIA, Clinical sensitivity, Clinical specificity, Donor screening, Early infection detection, HAV IgM, Hepatitis serology, HBsAg, Performance evaluation study, Seroconversion, Viral hepatitis elimination

Abstract

Background: Reliable screening for hepatitis serological markers is critical for early diagnosis, patient management, and transfusion safety, as acute infections are often asymptomatic. Thus, it is difficult to determine an infection by a clinical assessment alone, often remaining undiagnosed for years, leading to delayed treatment, disease progression and ongoing transmission. With this in mind, HBsAg, Anti-HCV, and HAV IgM serve as the foundational, first-line assays in the diagnostic and screening algorithms for these viruses. In support global viral hepatitis elimination goals, evaluations of commercially available hepatitis assays are critical to ensure they possess the sensitivity required for accurate laboratory detection, particularly during the early stages of infection when analyte levels are low.
Objectives: To perform an analytical performance evaluation of three commercial hepatitis tests, specifically assessing their capability to detect early seroconversion and low-level reactivity.

Methods: Anti-HCV (Antibody to Hepatitis C Virus), HBsAg (Hepatitis B Surface Antigen) and HAV IgM (IgM antibody to Hepatitis A virus) on three CLIA analytical systems were evaluated on three CLIA analytical systems: Alinity i (Abbott Laboratories Inc., Abbott Park, Illinois, USA), Maglumi X3 (Shenzhen New Industries Biomedical Engineering Co., Ltd., Shenzhen, China), and CL-1200i (Mindray Bio-Medical Electronics Co., Ltd., Shenzhen, China).

Study design: Analytical precision, analytical sensitivity, endogenous interferences of common metabolites, clinical sensitivity and clinical specificity were evaluated using reference materials, seroconversion panels and leftover patient samples, confirmed by nucleic acid amplification testing (PCR).

Results: Alinity i demonstrated the highest analytical sensitivity for HBsAg, remaining reactive at the lowest concentration evaluated (0.032 IU/mL), with complete (100%) clinical sensitivity across all assays. Maglumi X3 showed comparable HBsAg analytical sensitivity (0.043 IU/mL), very good clinical sensitivity (100% for Anti-HCV, 97.37% for HBsAg, 99% for HAV IgM), and high

specificity. CL-1200i exhibited reduced HBsAg analytical sensitivity (0.065 IU/mL), lower clinical sensitivity (91% for Anti-HCV, 84.87% for HBsAg, 92% for HAV IgM), with delayed seroconversion detection relative to Alinity i and Maglumi X3. All systems demonstrated high clinical specificity (>97.8%), although CL-1200i showed a higher rate (2.2% for Anti-HCV; 0.6% for HBsAg; 1.67% for HAV IgM) of false-reactivity relative to Alinity i (0.2% for Anti-HCV; 0.2% for HBsAg; 0.33% for HAV IgM) and Maglumi X3 (1.2% for Anti-HCV; 0.13% for HBsAg; 0.67% for HAV IgM).

Conclusions: Significant analytical and clinical performance differences exist among the three CLIA analytical systems evaluated in this study. This study provides analytical performance information that should aid medical laboratories when selecting a system for early infection detection and blood safety screening. Post-market evaluations of infectious disease immunoassays are increasingly critical under current IVDR requirements, particularly for early seroconversion detection where small analytical differences may translate into clinically relevant diagnostic delays and increased risk of missed infections.

Introduction

Hepatitis A virus (HAV), hepatitis B virus (HBV), and hepatitis C virus (HCV) infections remain major global public health concerns and continue to pose a significant risk to blood safety [1–3]. Despite substantial advances in screening strategies, transfusion-transmitted hepatitis remains a critical challenge, particularly in high throughput screening environments where serological testing represents the first line barrier to transmission [4,5]. Because acute hepatitis infections are frequently asymptomatic and cannot be reliably distinguished clinically, laboratory testing plays a central role in both individual patient management and population prevention strategies [4,6,7].

Blood donor screening programs rely on sensitive serological assays for the detection of hepatitis B surface antigen (HBsAg) and antibodies to HCV (anti-HCV) to prevent the transmission of chronic viral hepatitis [4,8]. Assay performance near clinical decision thresholds is of particular importance, as low level antigenemia or early seroconversion may occur during the early infection window or in cases of occult HBV or HCV infection [9]. Failure to detect such low reactivity samples may result in undetected infectious donations, with direct consequences for transfusion safety and public health [10].

Conversely, insufficient assay specificity may lead to false reactive results, unnecessary donor deferral, and increased confirmatory testing burden [8]. Detection of immunoglobulin M antibodies to HAV (HAV IgM) is central to the diagnosis of acute hepatitis A and to differentiating the infection from other causes of acute viral hepatitis [11]. Reliable HAV IgM testing is therefore essential for timely patient management, outbreak control, and epidemiological surveillance, particularly in emergency and outbreak driven testing settings [12,13].

Early and accurate laboratory detection of hepatitis A, B, and C infections is also fundamental to achieving global hepatitis control and elimination targets [1,2]. The World Health Organization (WHO) has emphasized that delayed or missed detection of hepatitis virus markers (particularly at low analyte concentrations) may contribute to ongoing transmission and delayed treatment initiation, thus undermining elimination efforts [2]. Consequently, analytical sensitivity, precision, and specificity at or near assay cutoff values represent clinically relevant performance characteristics for hepatitis serological assays used in routine diagnostic and blood screening practice [6,14].

Automated chemiluminescent immunoassay (CLIA) analytical systems are widely implemented for blood screening and infection detection in clinical laboratories due to their high throughput, automation, and standardized workflows[15]. However, differences in assay design, detection principles, and calibration approaches may lead to variability in analytical and clinical performance, particularly for low reactivity samples and early infection stages [16,17]. While numerous studies have evaluated individual hepatitis immunoassays, comparisons of widely used CLIA analytical systems across multiple hepatitis markers remain limited. In this context, real world performance evaluations are essential to characterize assay behaviour beyond manufacturer validation studies conducted under factory-controlled conditions [18,19].

The present study aims to provide a comprehensive analytical and clinical performance comparison of three automated CLIA analytical systems (Alinity i, Maglumi X3, and CL-1200i) commonly used for hepatitis serological screening. Specifically, we evaluated the analytical precision, analytical sensitivity, endogenous interference, and clinical sensitivity and specificity for Anti-HCV (Antibody to Hepatitis C Virus), HBsAg (Hepatitis B Surface Antigen), and HAV IgM (IgM antibody to Hepatitis A virus) under routine laboratory conditions.

Materials and Methods

Study design

This research was conducted as a head-to-head comparative evaluation of three automated chemiluminescent immunoassay systems. The study employed a cross-sectional experimental design to assess and compare the analytical and clinical performance of the Alinity i, Maglumi X3, and CL-1200i systems. To ensure a rigorous comparison, all systems were evaluated simultaneously using identical sets of reference materials, seroconversion panels, and patient serum samples under routine clinical laboratory conditions.

All measurements were performed by trained laboratory personnel using standard operating procedures, routine calibration, and manufacturer recommended settings to reflect a real-world diagnostic environment. Residual serum samples were prospectively collected after completion of routine diagnostic testing and were fully de-identified prior

to analysis. Samples were stored at -20 °C until testing.
The study was conducted at a private clinical laboratory in Bucharest, Romania, where all three analytical platforms were co-located. Due to institutional confidentiality policies, the specific name of the facility has been withheld. All procedures used fully de-identified residual serum samples collected after the completion of routine diagnostic testing. In accordance with the Clinical and Laboratory Standards Institute (CLSI) guidelines and the international standard ISO 15189:2022 [21], clinical laboratories must verify manufacturer performance claims using available relevant materials, including residual clinical specimens prior to the implementation of any system for patient reporting. Under the institutional and national regulatory standards in Romania, quality assurance evaluations using anonymized leftover specimens do not constitute interventional human subjects research. As such, the study was exempt from formal Ethics Committee approval and the requirement for informed consent. All data handling and procedures were conducted in strict adherence to the principles of the Declaration of Helsinki [20], CLSI guidelines, and in strict accordance with the instructions for use provided by the manufacturers.

Immunoassays and CLIA Devices

This study evaluated three automated CLIA analytical systems utilizing different chemiluminescent technologies. The Alinity i (Abbott Laboratories Inc., Abbott Park, Illinois, USA) and Maglumi X3 (Shenzhen New Industries Biomedical Engineering Co., Ltd., Shenzhen, China) utilize direct chemiluminescence (flash light production). These systems use luminescent labels attached directly to the antigen or antibody, resulting in fast reaction kinetics. In contrast, the CL-1200i (Mindray Bio-Medical Electronics Co., Ltd., Shenzhen, China) employs indirect chemiluminescence via glow light production. This system employs an enzyme-catalyzed reaction, typically utilizing an alkaline phosphatase label for a sustained light emission. Technical specifications for each CLIA system are summarized in Supplementary Table 1 and detailed information regarding the specific analytical and clinical parameters evaluated for Anti-HCV, HBsAg, and HAV IgM is presented in Table 1.

Table 1: The study protocol and materials used.

Assay	Precision	Analytical Sensitivity	Endogenous Interferents	Clinical Sensitivity	Clinical Specificity
AntiHCV	5 days (5 reps x 5 days) / Internal quality controls	N/A	Creatinine (15 mg/dL), Glucose (1000 mg/dL), Glycerol (18 mg/dL), Lactate (90 mg/dL), Pyruvate (5 mg/dL), Blood Urea Nitrogen (120 mg/dL), Uric Acid (25 mg/dL)	100 positive samples vs PCR	500 negative samples (50 pregnant females)
HBsAg		2nd WHO IS HBsAg (NIBSC 00/588)		250 samples (98 panel samples: low positive, occult, seroconversion; 152 positive patient samples vs PCR)	1500 negative samples
HAV IgM		N/A		100 positive samples	300 negative samples

N/A (Not applicable).

Analytical precision

The overall precision (within-lab precision) was defined as a combination of the within-run and between-run variances. The evaluation was conducted using manufacturer-provided internal quality controls according to the the Clinical and Laboratory Standards Institute CLSI EP15-A3 guidelines [21]. The protocol involved 5 replicates in a single run per day over 5 consecutive days, totaling 25 replicates for each quality control material.

Analytical sensitivity

Analytical sensitivity was defined as the lowest concentration of the WHO Reference Panel (NIBSC code 03/262) for HBsAg that yielded reactive results across replicate measurements on the analytical systems. Briefly, the panel consists of a series of fourfold dilutions (up to 1/256) of the 2nd WHO IS for HBsAg genotype A2, subtype group adw2 (NIBSC 00/588) [22]. The reference panel is comprised of four positive samples (panel members A-D) and one negative control sample (panel member E; normal re-calcified plasma).
Initial sensitivity testing was performed using panel members A-D and E (the negative control). To further characterize assay performance at low HBsAg concentrations near the detection limit, additional dilutions of panel member D (1/256 of IS 00/588) were prepared, generating three supplementary concentrations (D1-D3). Panel member E was used as matrix material to prepare additional dilutions of panel member D (1/256 of IS 00/588). Assigned concentrations for each dilution were calculated based on the assigned unitage of the 2nd WHO International Standard (33 IU/mL). These extended dilutions were included to allow comparative assessment of analytical sensitivity beyond the standard panel range. The final extended dilution was determined when two out of the three CLIA assay systems failed to detect the diluted sample. All panel members

and additional dilutions were tested on the same day on each analyzer, with five replicate measurements per concentration. Analytical sensitivity in this study was reported as a comparative performance characteristic and does not represent a formal limit of detection (LoD) determination.

Endogenous interferences

Analytical interference for each assay was evaluated with seven common endogenous metabolites expected in the patient population at high concentrations using the ASSURANCE™ Interference Test Kit INT-06 (Sun Diagnostics, LLC, New Gloucester, Maine, USA). This kit contained the following metabolites, with targeted concentrations shown in parentheses: creatinine (15 mg/dL), glucose (1000 mg/dL), glycerol (18 mg/dL), lactate (90 mg/dL), pyruvate (5 mg/dL), blood urea nitrogen (120 mg/dL), uric acid (25 mg/dL). Leftover patient serum samples without any visible lipemia, haemolysis or icterus were used to prepare negative, low positive and high positive analyte pools. Serum pools were tested before spiking to establish the baseline measurand concentrations, then test pools were prepared by adding the interferent material to bring the test pool to the target interferent concentration (Table 1). The required amount of interferent materials were calculated using the Assurance™ Excel Worksheets [23]. For each test pool, a matching control pool with the same volume of solvent (purified water) as used in preparation of the corresponding interferent pool was prepared. Same day measurements for each potential interferent were performed by measuring 6 replicates per run for each sample pool on all the analyzers. Furthermore, to minimize experimental error, the replicates were analyzed in alternating pattern between control pools and test pools. Because the evaluated immunoassays generate quantitative results that are interpreted qualitatively, interference acceptability was assessed relative to the assay cutoff rather than fixed percentage bias limits. Specifically, the maximum allowable difference (dmax) was established relative to the medical decision point (immunoassay cutoff). For negative samples pool, dmax was defined as the maximum allowable signal shift that would not cause a true negative result to exceed the positivity cutoff. For the positives and the weak positives pools, dmax was defined as the maximum allowable signal shift that would not reduce a true positive result below the cutoff.

Clinical sensitivity

Clinical sensitivity evaluation was conducted separately for each assay, using a combination of leftover patient samples and commercially available seroconversion and sensitivity panels (Table 1). Anti-HCV clinical sensitivity was assessed using 100 positive leftover patient samples, confirmed by HCV RNA PCR. HAV IgM clinical sensitivity was evaluated using 100 positive leftover patient samples, classified as positive based on repeated reactivity, defined as duplicate reactive results obtained on at least two of the three evaluated analyzers (as per manufacturer instructions). HBsAg clinical sensitivity was evaluated by testing a total of 250 samples (comprising 98 panel samples and 152 positive leftover patient samples). Furthermore, the positive leftover samples were also confirmed positive by HBV DNA PCR for validation.
Seroconversion performance was assessed using six commercially available HBsAg seroconversion panels 0605-0038; 0605-0039; 0605-0041; 0605-0059 (SeraCare, LCG Group, LCG Limited, Guildford, UK [24]), and 67303; 65449 (ZeptoMetrix LLC, Franklin, Massachusetts, USA [25]), consisting of sequential plasma samples collected from individual donors during HBV seroconversion. Additional sensitivity assessment was performed using AccuSet Low Titre HBsAg Performance Panel 0805-0417; AccuSpan Sensitivity Panel HBsAg Subtype 2410-0161 (SeraCare, LCG Group, LCG Limited, Guildford, UK). The low titre performance panel consisted of undiluted HBV positive plasma samples with reactivities close to the cutoff signal of these commercial HBsAg assays. The sensitivity panel consists of two series of dilutions for HBsAg subtypes ad and ay, designed to evaluate low end assay sensitivity. The panel manufacturer tested the WHO HBsAg subtype adw2, genotype A Reference Panel (NIBSC code 03/262) in the same run as the AccuSpan HBsAg Sensitivity Panel members [24].

Clinical specificity

Clinical specificity evaluation was conducted separately for each assay using leftover negative patient samples. The varying sample sizes across the evaluated markers reflect the practical challenges of prospectively acquiring sufficient volumes of adequately characterized uncontaminated leftover diagnostic specimens during the study period. Based on this sample availability, evaluations were performed as follows: Anti-HCV using 500 negative patient samples (including 50 samples obtained from pregnant women); HBsAg using 1,500 negative patient samples; and HAV IgM using 300 negative patient samples. Reactive (positive) results detected for Anti-HCV, HBsAg, and HAV IgM assays were further investigated using confirmatory nucleic acid amplification testing together with clinical information (where available) to determine whether an active infection was present or if the result was falsely reactive.

Statistical analysis

Analyses were performed to evaluate the comparative performance of the three analytical systems. Precision was quantified using the coefficient of variation (CV%) and analyzed via analysis of variance (ANOVA, using the day of the run as random factor variable) to determine within run and total within laboratory signal stability. For clinical sensitivity and specificity, point estimates were calculated as proportions of correctly identified samples, with 95% confidence intervals (CI) determined using the exact binomial Clopper-Pearson method.

Table 2: Repeatability (CV%) results compared to the manufacturer's claimed values.

Assay	Control sample	N	Alinity i		Maglumi X3		CL1200i	
			Claimed	Measured	Claimed	Measured	Claimed	Measured
AntiHCV	Negative Control	25	7.00%	5.35%	13.24%	5.12%	8.98%	9.83%
	Positive Control	25	3.20%	1.48%	4.22%	3.89%	5.86%	4.82%
HBsAg	Negative Control	25	N/A	7.08%	N/A	14.69%	8.00%	10.70%
	Positive Control	25	3.20%	2.07%	4.71%	3.54%	4.49%	7.49%
	Positive Control 2	25	N/A	N/A	4.66%	3.23%	N/A	N/A
HAV IgM	Negative Control	25	7.00%	6.80%	N/A	8.08%	3.41%	4.17%
	Positive Control	25	6.30%	3.39%	1.96%	3.90%	2.25%	4.55%

N/A (Not available). Green shading indicates measured CV% values meeting or exceeding manufacturer performance claims (≤ claimed CV%). Red shading indicates measured CV% values higher than the manufacturer's claimed precision.

Regarding endogenous interference studies, patient serum samples were pooled to create standardized negative, low positive, and high positive analyte levels. Pooling was necessary to provide sufficient sample volume for multi replicate testing and to establish a stable baseline measurand concentration before the addition of interferent materials. The acceptability of interference was assessed using the maximum allowable difference (dmax) relative to the medical decision point (assay cutoff). Finally, the robustness of each system against false reactivity was assessed by measuring the statistical separation between the negative sample distribution and the assay cutoff, expressed in standard deviations (SD).

Results

Analytical precision results

The analytical precision of the three evaluated CLIA assay systems was determined by assessing both within-run repeatability and total within-laboratory reproducibility across all three hepatitis markers (Anti-HCV, HBsAg, and HAV IgM). Precision was computed using the coefficient of variation (CV%) across multiple replicates of multi-level control materials. Statistical parameters for precision are summarized in Table 2 (exhaustive statistical parameters shown in Supplementary Table 2).

The Alinity i platform exhibited the best relative precision metrics throughout the evaluation. For this system, the measured CV values consistently fell well below the manufacturer's claimed limits (see Table 2). This robust signal stability was uniformly maintained across all evaluated assays and across both negative and positive control levels.
The Snibe Maglumi X3 and Mindray CL-1200i systems exhibited higher CV percentages compared to the Alinity i [26]. Specifically, higher imprecision (elevated CV percentages, see Table 2) was observed on both systems when evaluating negative control materials at concentrations near the assay cutoff thresholds.
At the positive control levels, both the Alinity i and Maglumi X3 systems demonstrated CVs below their claimed limits, whereas the CL-1200i exhibited comparatively higher CVs. For these assay specific controls, both the Alinity i and Maglumi X3 systems demonstrated good signal stability (see Supplementary Table 2), whereas the CL-1200i exhibited comparatively higher (though still acceptable) precision and signal fluctuation.

Analytical sensitivity results for HBsAg

Analytical sensitivity results for the HBsAg assays are summarized in Table 3. All three systems measured reactive results across the panel members A-D, corresponding to HBsAg concentrations from 8.25 IU/mL to 0.13 IU/mL. At lower concentrations generated by additional dilutions of panel member D, differences in analytical sensitivity were observed between the evaluated systems. Alinity i remained reactive across all tested dilutions, including the lowest concentration evaluated (0.032 IU/mL; D3), with consistent reactivity observed across all five replicate measurements. Maglumi X3 remained reactive at 0.065 IU/mL (D1) and at 0.043 IU/mL (D2). At the 0.043 IU/mL dilution, three out of five replicate measurements were above the assay cutoff (≥0.05 IU/mL), while two replicates were below the cutoff. According to the manufacturer's instructions for use, this pattern was classified as reactive. At the lowest concentration tested (0.032 IU/mL; D3), all replicate measurements were non-reactive. CL-1200i remained reactive at 0.065 IU/mL (D1). At 0.043 IU/mL (D2), one out of five replicate measurements exceeded the assay cutoff, while the remaining replicates were below the cutoff; according to the manufacturer's instructions for use, the result was classified as non-reactive. At 0.032 IU/mL (D3), all replicate measurements were non-reactive.

Table 3: HBsAg analytical sensitivity comparison. WHO Reference Reagent HBsAg subtype adw2, genotype A, WHO Reference Panel (NIBSC 03/262): Dilutions of the 2nd WHO International Standard for HBsAg genotype A2 subtype group adw2 (NIBSC 00/588) Assigned unitage: 33 IU/mL.

Panel members	Dilution unitage (IU/mL)	Alinity i (S/CO)	Maglumi X3 (IU/mL)	CL1200i (IU/mL)
A (1/4 IS)	8.25	929.282	10.231	9.291
B (1/16 IS)	2.06	331.676	2.775	3.549
C (1/64 IS)	0.52	126.550	0.400	0.793
D (1/256 IS)	0.13	32.910	0.198	0.106
D1 (1/2 D)	0.065	12.978	0.056	0.073
D2 (1/3 D)	0.043	8.782	0.051	0.048*
D3 (1/4 D)	0.032	6.618	0.023	0.039
E (Negative)	0	0.200	0.023	0.011

*Positive results: green highlight, red line: detection limit; *1 reactive result out 5 replicates tested, the result is considered negative. According to the instructions for use: Samples with HBsAg concentration ≥0.05 IU/mL are considered Initially Reactive. Initially Reactive samples must be retested in duplicate, if both results are less than 0.05 IU/mL, the sample is non-reactive; no further test is needed for the patient.*

Endogenous interference evaluation results

No clinically significant interference was observed from common endogenous substances tested according to CLSI EP07 for all three CLIA systems and infectious disease immunoassays (Data not shown). Differences were observed (in some cases exceeding ±10% relative difference between control and test pools) for some runs but this did not alter the qualitative interpretation of results (positive/negative), as no result crossed the assay cutoff. Overall, Snibe Maglumi X3 immunoassays demonstrated slightly higher signal variation compared to Alinity i and CL-1200i. Notably, for Anti-HCV, Maglumi X3 showed significant concentration differences of 10.1-25.2% for creatinine (15 mg/dL) across negative, low positive, and positive pools, while CL-1200i showed differences of 5.3-10.4% for uric acid (25 mg/dL). For HBsAg, Maglumi X3 showed differences of 5.1-17.6% for lactate (90 mg/dL), while CL-1200i showed differences of 4.3-10.7% for lactate (90 mg/dL). For all other endogenous metabolites, signal differences remained within ±10%. Accordingly, interference effects exceeding ±10% were considered analytically but not clinically significant for qualitative interpretation [27].

Clinical sensitivity evaluation results

Clinical sensitivity evaluation results for Anti-HCV

Alinity i and Maglumi X3 demonstrated a clinical sensitivity of 100%, correctly identifying all 100 positive samples, with no false negative results observed (Table 4, with results distribution shown in Supplementary Figure 1A). In contrast, CL-1200i correctly classified 91 of 100 positive samples, corresponding to a clinical sensitivity of 91.0% (95% CI: 83.77-95.19%), with 9 false negative results. The majority of reactive samples on all systems were distributed well above the assay cutoff (S/Co, AU/mL, or COI = 1). For CL-1200i, false negative results were predominantly associated with samples exhibiting low signal values close to the cutoff (on the other systems), indicating reduced sensitivity in the low reactive range compared to Alinity i and Maglumi X3.

Table 4: Clinical sensitivity results.

Assay	Analyzer	Cutoff	Unit	N	Negative	Positive	Mean	Sensitivity	Sensitivity 95%CI
AntiHCV	Alinity i	1	S/CO	100	0	100	9.85	100%	(96.30% / 100.00%)
	Maglumi X3	1	AU/mL	100	0	100	52.37	100%	(96.30% / 100.00%)
	CL1200i	1	COI	100	9	91	42.07	91%	(83.77% / 95.19%)
HBsAg	Alinity i	1	S/CO	152	0	152	66.76	100%	(97.54% / 100.00%)
	Maglumi X3	0.05	IU/mL	152	4	148	1.17	97.37%	(93.43% / 98.97%)
	CL1200i	0.05	IU/mL	152	23	129	0.07	84.87%	(78.32% / 89.70%)
HAV IgM	Alinity i	0.8	S/CO	100	0	100	7.38	100%	(96.30% / 100.00%)
	Maglumi X3	1	AU/mL	100	1	99	6.49	99%	(94.55% / 99.82%)
	CL1200i	1	COI	100	8	92	5.74	92%	(85.00% / 95.89%)

Clinical sensitivity evaluation results for HBsAg

Alinity i demonstrated the highest clinical sensitivity, correctly identifying all 152 positive samples (100%; 95% CI: 97.54-100.00%), with no false negative results observed. Maglumi X3 identified 148 of 152 positive samples, corresponding to a clinical sensitivity of 97.37% (95% CI: 93.43-98.97%), while CL-1200i correctly classified 129 of 152 positive samples, yielding a clinical sensitivity of 84.87% (95% CI: 78.32-89.70%) (see Table 4). A higher proportion of low signal samples near the cutoff was observed for Maglumi X3 and CL-1200i, consistent with the increased number of false negative results on these systems (Supplementary Figure 1B).

Evaluation using the HBsAg Low Titer Performance Panel and the HBsAg Sensitivity Panel further highlighted differences among the three systems (Table 5A). Alinity i correctly identified all panel samples, whereas Maglumi X3 failed to detect 3 low titer positive samples and CL-1200i identified correctly only 12 of the 20 low titer positive samples (Table 5B). In the HBsAg Sensitivity Panel 2410-0161, all HBV subtypes were detected by Alinity i and Maglumi X3, while one subtype (ay, panel member 14) was not detected by CL-1200i. Seroconversion panel testing demonstrated earlier HBV detection for Alinity i compared to the other systems (see Table 5A). Across the six seroconversion panels, Alinity i showed the shortest mean time to first HBsAg detection (10.7 days), followed by Maglumi X3 (21.2 days), while CL-1200i exhibited the longest mean detection time (28.5 days). In several panels, CL-1200i did not detect HBsAg within the observation period (Supplementary Figure 2).

Table 5: Days until detection and Low Titre Clinical Sensitivity for HBsAg.

No. Crt.	Seroconversion Panel ID	Alinity i	Maglumi X3	CL-1200i
1	0605-0041	0	11	14
2	0605-0039	0	9	13
3	0605-0038	0	5	7
4	0605-0059	2	18‡	>28 (ND)*
5	67303	62	72	79
6	65449	0	12	26
Mean number of days until detection		**10.7**	**21.2**	**28.5**
Number of positive results (out of total 104 test results)		**80 (76.9%)**	**52 (50%)**	**32 (30.7%)**

No. Crt.	Seroconversion Panel ID	Alinity i	Maglumi X3	CL-1200i
(B)				
1	0805-0417	20/20	17/20	12/20
2	2410-0161	15/15	15/15	14/15
Number of correctly identified results (out of total 70 duplicate test results)		**70 (100%)**	**64 (91.4%)**	**52 (74.3%)**

(A) Days until 1st HBsAg detection (Seroconversion Panels). No detection is denoted with ND. Asterisk (*) indicates that no signal was detected. Cross (‡) denotes that the Maglumi X3 device had intermittent detection (on days 18 and 28 – see Supplementary Figure 2)
(B) Clinical sensitivity evaluation for HBsAg determined using the Low Titre Performance Panel and the HBsAg Sensitivity Panel (Correctly Identified Samples).

Clinical sensitivity evaluation results for HAV IgM assay

Alinity i demonstrated a clinical sensitivity of 100%, correctly identifying all 100 positive samples, with no false negative results observed (see Table 4). Maglumi X3 correctly identified 99 of 100 positive samples, corresponding to a clinical sensitivity of 99.0% (95% CI: 94.55-99.82%), while CL-1200i identified 92 of 100 positive samples, yielding a clinical sensitivity of 92.0% (95% CI: 85.00-95.89%). Of the CLIA systems tested, CL-1200i exhibited a higher proportion of low signal results close to the cutoff, consistent with the increased number of false negative results compared to Alinity i and Maglumi X3 (Supplementary Figure 1C).

Clinical specificity evaluation results

Clinical specificity evaluation results for Anti-HCV

Clinical specificity results for the Anti-HCV assays (Table 6) reveal that from a total of 501 leftover patient samples tested, 19 samples were found reactive on at least one analyzer. One sample was confirmed positive by HCV RNA PCR testing (HCV viral load of 5,210 IU/mL) and was excluded from the specificity analysis, this sample was correctly identified as positive on Alinity i (mean result of 4.030 S/Co) and Maglumi X3 (mean result of 16.824 AU/mL), but non-reactive on CL-1200i.The remaining 18 samples were confirmed negative and were considered false reactive results. Clinical specificity calculations were therefore performed using the remaining 500 negative samples. Alinity i demonstrated the highest clinical specificity for the Anti-HCV assay, with a specificity of 99.80% (95% CI: 98.89 to 99.99%), corresponding to one false reactive result among the 500 negative samples. Maglumi X3 exhibited 98.80% (95% CI: 97.41 to 99.56%), with six false reactive results, while CL-1200i showed the lowest performance at 97.80% (95% CI: 96.10-98.90%), with eleven false reactive results.

Table 6: Clinical specificity.

Assay	Analyzer	Cutoff	Unit	N	Negative	Positive	Mean	Specificity	Sensitivity 95%CI
AntiHCV	Alinity i	1	S/CO	500	499	1	0.13	99.80%	(98.89% / 99.99%)
	Maglumi X3	1	AU/mL	500	494	6	0.36	98.80%	(97.41% / 99.56%)
	CL1200i	1	COI	500	489	11	0.20	97.80%	(96.10% / 98.90%)
HBsAg	Alinity i	1	S/CO	1500	1497	3	0.32	99.80%	(99.42% / 99.96%)
	Maglumi X3	0.05	IU/mL	1500	1498	2	0.03	99.87%	(99.52% / 99.98%)
	CL1200i	0.05	IU/mL	1500	1491	9	0.02	99.40%	(98.86% / 99.73%)
HAV IgM	Alinity i	0.8	S/CO	300	299	1	0.13	99.67%	(98.16% / 99.99%)
	Maglumi X3	1	AU/mL	300	298	2	0.47	99.33%	(97.61% / 99.92%)
	CL1200i	1	IU/mL	300	295	5	0.55	98.33%	(96.15% / 99.46%)

The vast majority of negative results (including the 50 samples from pregnant women) clustered well below the assay cutoffs across all three CLIA systems (Supplementary Figure 3A). The Alinity i demonstrated the greatest statistical separation between the negative sample distribution and the cutoff threshold (distance to cutoff: 12.33 standard deviations – SD), whereas the Maglumi X3 and CL-1200i exhibited narrower margins (1.15 SD for both). While false-reactive results were predominantly observed close to the assay cutoffs, isolated samples with higher reactive values were recorded on the CL-1200i and Maglumi X3.

Among the subgroup of 50 pregnant patients, one false-reactive result was observed on the Alinity i. This sample tested negative on both the Maglumi X3 and CL-1200i. The Maglumi X3 produced a false-reactive result from a different sample within this cohort. The CL-1200i produced zero false-reactive results among the pregnant patient samples.

Clinical specificity evaluation results for HBsAg

Clinical specificity results for the HBsAg assays (Table 6) reveal that from a total of 1,500 leftover patient samples tested, 13 samples were found reactive on at least one analyzer (one sample was false positive on both Alinity i and Maglumi X3). All reactive samples were confirmed negative by HBV DNA PCR testing and were considered false positive results. Alinity i demonstrated a clinical specificity of 99.80% (95% CI: 99.42 to 99.96%), with three false reactive results, while Snibe Maglumi X3 showed 99.87% (95% CI: 99.52 to 99.98%), corresponding to two false reactive results. Mindray CL-1200i exhibited 99.40% (95% CI: 98.86 to 99.73%), with 9 false reactive results among the 1,500 negative samples.

The majority of negative samples across all assays clustered well below their respective cutoffs (Supplementary Figure 3B). Reflecting the false-positive rates, the systems demonstrated distinct statistical separation between the negative sample distributions and their respective cutoff thresholds. The Alinity i showed the widest margin (distance to cutoff: 5.83 SD), followed by the Maglumi X3 (3.18 SD). In contrast, the CL-1200i exhibited a notably narrower margin (0.53 SD), which aligns with its higher incidence of false reactivity.

Clinical specificity evaluation results for HAV IgM assay

Clinical specificity results for the HAV IgM assays (Table 6) show that from a total of 300 samples tested, 6 were found reactive on at least one analyzer. All reactive samples were confirmed negative by HAV RNA PCR testing and were considered false positive results. One of the reactive samples was identified as positive on all three analyzers (mean result of 2.237 S/Co on Alinity i; 3.235 AU/mL on Maglumi X3; 5.130 COI on CL-1200i), however, HAV RNA was not detected, and no clinical evidence of acute HAV infection was available, therefore, this sample was classified as false reactive. Alinity i demonstrated a clinical specificity of 99.67% (95% CI: 98.16 to 99.99%), corresponding to one false positive result among the 300 negative samples. Maglumi X3 showed 99.33% (95% CI: 97.61 to 99.92%), with two false positive results, while CL-1200i exhibited 98.33% (95% CI: 96.15 to 99.46%), corresponding to five false positive results.

The majority of negative samples across all three systems clustered well below their respective assay cutoffs (Supplementary Figure 3C). Consistent with the observed false-reactive profiles, the systems displayed distinct degrees of statistical separation between their negative sample distributions and the cutoff thresholds. The Alinity i demonstrated the widest margin (distance to cutoff: 5.34 SD), followed by the Maglumi X3 (2.30 SD). In contrast, the CL-1200i exhibited the narrowest separation (0.68 SD), with its false-reactive results primarily occurring at low signal values just above this threshold.

Discussion

Our evaluation reveals that while all tested analytical systems are suitable for routine screening, their distinct performance profiles at the low reactive end have significant implications for clinical decision making and public health outcomes. These differences are most pronounced in the detection of nascent infections and low level antigenemia, suggesting that the selection of an analytical system must be guided by the specific diagnostic requirements of the patient population and the screening environment rather than operational metrics alone.

Analytical precision

All systems demonstrated acceptable analytical precision for routine diagnostic use. Alinity i consistently achieved precision values well below manufacturer claimed limits across all assays reflecting high within and total lab precision [28–30]. Maglumi X3 and CL-1200i also demonstrated acceptable precision performance although higher variability was observed, particularly for negative control materials near the cutoff values. The increased imprecision observed with the negative or near cutoff level controls across all systems does not affect clinical interpretation, as results remain below reactive thresholds. Importantly, the precision observed with positive control levels, which is clinically relevant for result confirmation and patient monitoring, was robust for all systems. Instances in which manufacturer claimed precision specifications were not fully met likely reflect the controlled manufacturer validation environments [31,32].

Analytical sensitivity and low level HBsAg detection

While the WHO guidance emphasizes the importance of using highly sensitive HBsAg assays, it does not prescribe a numeric requirement for analytical sensitivity [4]. However, we demonstrate that substantial differences exist among the three CLIA systems, particularly at low antigen concentrations near the assay cutoffs. The Alinity i HBsAg Next Qualitative assay demonstrated the the highest analytical sensitivity for low level antigen detection, followed by the Maglumi X3, while the CL 1200i showed a narrower range of reactivity. These findings

are broadly consistent with manufacturer reported analytical sensitivity claims [28,30].

From a clinical perspective, enhanced analytical sensitivity for HBsAg detection is particularly important for blood donor screening, early infection detection, and the identification of low level or occult HBV infections [9]. In these contexts, small analytical differences at low antigen concentrations may translate into delayed or even failed detection during early seroconversion stages or for samples with low level antigen concentration. This has important implications for transfusion safety and clinical management [34]. Furthermore, in high throughput screening settings, the differences in analytical sensitivity become crucial given that serological assays serve as the primary barrier prior to nucleic acid testing [4,5].

Endogenous interferences

None of the tested substances produced clinically significant interference across the three evaluated CLIA systems. Although statistically significant variations were observed under selected conditions, these effects did not alter qualitative result interpretation, as no sample crossed the assay cutoff. In accordance with CLSI EP07 guidance, interference effects were interpreted based on their potential impact on clinical decision rather than statistical significance alone [27]. For qualitative immunoassays, analytical signal variations exceeding predefined relative difference thresholds may be acceptable provided that reactive versus non-reactive classification remains unchanged [27,35]. In this study, all observed interference effects, including those exceeding ±10% relative difference, were therefore considered analytically but not clinically significant. Among the evaluated systems, Snibe Maglumi X3 demonstrated slightly greater signal variability in the presence of selected endogenous substances, particularly creatinine and lactate, compared with Alinity i and CL-1200i. However, these variations did not compromise result classification on any platform. Overall, the findings confirm adequate resistance to clinically relevant endogenous interference for all three systems, supporting their suitability for routine diagnostic and screening applications [27,35].

Clinical sensitivity and early seroconversion performance

Across all three viral hepatitis markers evaluated, Alinity i consistently demonstrated better clinical sensitivity, detecting all confirmed positive cases across the assay spectrum. Maglumi X3 provided a slightly lower, overall detection profile. In contrast, while performing within clinically acceptable margins for general routine screening, CL-1200i exhibited a pronounced reduction in identifying true positive samples. Notably, the false-negative results observed on the CL-1200i (and to a lesser extent on Maglumi X3) were predominantly associated with samples exhibiting low signal values near the assay cutoffs on the other systems. This suggests a limitation in detecting weakly reactive samples which is important for the identification of early-stage infections, low-titer presentations, and immunologically atypical patients.

While this analytical limitation is consistent across the assays, the clinical and public health consequences vary critically by the target antigen. For Anti-HCV, failure to detect weakly reactive samples creates a meaningful diagnostic gap. Missing these infections delays confirmatory RNA testing, treatment initiation, and linkage to care [4,36]. This is especially concerning in prenatal screening, where altered maternal antibody responses can occur, and undetected infection prevents appropriate follow-up, counselling, and postnatal monitoring of exposed infants [37].

For HBsAg, the risk of low-end sensitivity limitations is most vividly illustrated during early seroconversion. By substantially narrowing the diagnostic window, Alinity i detected nascent HBV infections considerably earlier and with greater frequency than the other systems. Conversely, CL-1200i missed a substantial proportion of low-titer samples and exhibited markedly delayed (or entirely absent) detection across full seroconversion series. This poses a direct threat to high-throughput blood donor screening, where identifying window-period infections is the primary barrier against onward transmission [8].

Lastly, for HAV IgM, because the clinical presentation of acute viral hepatitis is indistinguishable across different viral etiologies, rapid serological confirmation is paramount [11]. Missing low-level HAV IgM reactivity delays essential infection control measures and outbreak management. Furthermore, while HAV does not establish chronic infection, it can lead to fatal fulminant acute liver failure [38]. Relying on robust front-line serology is essential, as nucleic acid testing (while capable of identifying early infections) is often too costly and specialized for routine acute diagnosis [11,38].

Ultimately, these findings underscore that the clinical sensitivity of CLIA systems is a paramount consideration in assay selection, highlighting the clinical utility of a stratified, fit-for-purpose testing strategy [7]. Cost-effective systems that perform within acceptable clinical margins may be appropriately utilized for the routine monitoring of known positive patients, where analyte titers are typically well-established and the risk of missing a nascent infection is absent. Conversely, deploying systems with maximal analytical sensitivity (such as Alinity i) is imperative for initial screening in high-risk scenarios (blood donation, diagnostically complex patients, and early infection detection) where minimizing the diagnostic window period is critical [4].

In particular, highly sensitive screening in pregnant women is vital to detect low-level reactivity and ensure appropriate clinical interventions that prevent mother-to-child transmission [39]. While systems with maximal sensitivity, such as Alinity i, are optimal for preventing missed infections in this high-risk group, our evaluation revealed distinct specificity profiles within this demographic. In our cohort of pregnant patients, the CL-1200i platform demonstrated superior specificity for Anti-HCV screening by yielding zero false-reactive results. Conversely, the isolated false-reactive Anti-HCV results

observed on both the Alinity i and Maglumi X3 systems originated directly from this pregnant cohort, highlighting a well-documented clinical trade-off where ultra-high analytical sensitivity may occasionally capture pregnancy-induced immunological cross-reactivity. Therefore, the strategic deployment of highly sensitive screening assays is a fundamental requisite for advancing the WHO's goal of eliminating viral hepatitis as a public health threat by 2030 [1,2].

Clinical specificity and false-reactive

All evaluated systems demonstrated acceptable clinical specificity (>97.8%), confirming their overall diagnostic reliability for routine screening. However, distinct performance differences emerged among the systems. Alinity i and Maglumi X3 consistently exhibited the lowest false-positive rates across the evaluated markers. In contrast, while CL-1200i performed well within acceptable clinical parameters, it demonstrated a greater incidence of borderline reactivity and a comparatively higher proportion of false-reactive results, particularly in the Anti-HCV and HAV IgM assays.

These distinct performance profiles are directly explained by the assays' statistical separation between their negative sample distributions and their respective cutoff thresholds. Across all three viral markers, Alinity i maintained the widest statistical margin (measured as standard deviations – SD – to the cutoff), indicating that its baseline negative population clusters safely away from the reactive threshold. This robust margin makes the system highly resistant to minor biological noise and non-specific cross-reactivity. Conversely, the CL-1200i consistently exhibited the narrowest SD margins. When the baseline negative distribution sits in such close proximity to the cutoff, even minor assay signal fluctuations or low-level endogenous interference can push a true-negative sample over the threshold. This dynamic directly accounts for the higher incidence of borderline and false-reactive results observed on the Mindray platform. Lastly, Maglumi X3 maintained an intermediate separation that proved sufficient to sustain high specificity.

A significant difference regarding Anti-HCV testing in pregnant patients was observed. The false-reactive results identified by Alinity i and Maglumi X3 occurred within this specific cohort, likely reflecting the known immunological alterations characteristic of pregnancy [40]. Specifically, pregnancy induces the production of cross-reactive maternal alloantibodies directed against fetal antigens, as well as transient increases in heterophile antibodies. These circulating proteins can non-specifically bind to the assay's capture or detection reagents, generating spurious reactive signals despite the absence of true viral specific antibodies.

From an operational and public health perspective, the observed differences in false-positive rates, although small on a percentage basis, can have substantial practical implications for high-volume screening laboratories. While false-reactive results do not carry the immediate transmission risks of a missed infection, their downstream effects are costly and disruptive. Suboptimal specificity inevitably triggers a cascade of unnecessary confirmatory PCR testing, unwarranted patient anxiety, and unjustified blood donor deferrals [41,42]. Consequently, prioritizing high-specificity systems is essential for minimizing the clinical, financial, and operational burdens inherent to large-scale screening programs.

Study limitations and future perspectives

To our knowledge, this is the first study to provide a comprehensive, head-to-head performance comparison of the Abbott Alinity i, Snibe Maglumi X3, and Mindray CL-1200i CLIA systems across multiple viral hepatitis markers. This comparative evaluation fills a critical gap in the literature, as real-world assessments are essential to guide laboratory assay selection and characterize system behavior in a medical laboratory.

This study has several limitations. Because the evaluation was conducted retrospectively at a single center, the generalizability of the findings to other laboratory environments, workflows, or analyzer configurations may be limited. Furthermore, the assessment of viral genotype diversity was inherently restricted to the specific WHO reference materials and commercial panels available for inclusion. Long-term reagent stability and lot-to-lot variability were also not evaluated. Consequently, the results represent a snapshot of assay performance under the specific testing conditions, reagent lots, analyzer configurations, and software versions utilized during the study period. As such, performance characteristics may vary across different laboratories and over time.

Nevertheless, this foundational assessment establishes a robust baseline for these widely used screening systems. Future prospective studies incorporating multicenter evaluations, broader geographic datasets, and a wider array of clinical materials will be invaluable to further support, validate, and extend the findings of this comparative assessment.

Conclusion

Our study revealed significant performance disparities among the three CLIA systems. Based on the cumulative analytical and clinical data, this study establishes the comparative performance profiles for hepatitis serological screening: Alinity i demonstrated superior reliability over the Maglumi X3, which in turn outperformed the CL 1200i.

The Alinity i system provided the highest level of clinical sensitivity at 100% across all evaluated assays and achieved the most rapid detection of early HBV seroconversion. While the Maglumi X3 demonstrated high overall clinical sensitivity and specificity, it exhibited slightly lower sensitivity for low titer HBsAg samples and minor delays in seroconversion detection compared to the Alinity i. In contrast, the CL 1200i system exhibited a critical diagnostic defect, exhibiting a reduced clinical sensitivity of 84.87% for HBsAg and - delayed detection during early infection stages. These analytical

limitations directly translated into longer diagnostic window periods and a higher incidence of misclassified results compared to the other systems. For high throughput environments where early detection is the primary barrier to disease transmission, the use of analytical systems with limited low-end sensitivity poses a substantial risk to blood safety and public health. Consequently, laboratory selection must prioritize the categorical evidence of performance to ensure effective clinical intervention and actively support global hepatitis elimination targets.

Conflict of interest statement

D.G.C., S.D., and A.J.N. are affiliated to Dig Worldwide Ltd. Dig Worldwide Ltd was involved in the study design, data collection, analysis, and the decision to submit the manuscript for publication. D.G.C. is a contracted specialist working with Dig Worldwide Ltd, S.D. is an employee of, and A.J.N. is a shareholder of Dig Worldwide Ltd. Abbott Laboratories had visibility towards the data and the manuscript before submission without altering the results or their interpretation. The final publication decision remained independently with the authors. The authors declare no other competing interests.

Funding statement

This study was supported by Dig Worldwide Ltd., a commercial consultancy firm in the life sciences sector, utilizing project funding from Abbott Laboratories. This research was an initiative of Dig Worldwide Ltd. and the results presented in this study were not influenced or biased in any way by CLIA system manufacturers or other second or third parties.

Data availability statement

The analyzed data that support the findings of this study are available on request from the corresponding author, by emailing: dgc@dig-worldwide.com.

CRediT authorship contribution statement

Author Contributions: Conceptualization, D.C., S.D., A.N.; validation D.C., S.D., A.N.; investigation D.C., formal analysis S.D., D.C., A.N.; data curation D.C.; writing - original draft D.C.; writing - review & editing D.C., S.D., A.N.; visualization D.C., project administration D.C., S.D., A.N.; supervision A.N.; funding acquisition A.N.;

References

1. World Health Organization. Global health sector strategies on HIV, viral hepatitis and sexually transmitted infections for the period 2022-2030 https://www.who.int/publications/i/item/9789240053779 (accessed April 2, 2026).
2. CDC Reports Latest National Data on Syphilis in Newborns and Sexually Transmitted Infections (STIs) | CDC Newsroom n.d. https://www.cdc.gov/media/releases/2025/2025-cdc-reports-latest-national-data-on-syphilis-in-newborns-and-sexually-transmitted-infections-stis.html (accessed April 2, 2026).
3. World Health Organization. Global health sector strategies on, respectively, HIV, viral hepatitis and sexually transmitted infections for the period 2022-2030. GHSS 2022-2030 2022.
4. UNAIDS 2025 Global AIDS Update — AIDS, Crisis and the Power to Transform | UNAIDS n.d. https://www.unaids.org/en/resources/documents/2025/2025-global-aids-update-summary (accessed April 2, 2026).
5. Ghanem KG, Ram S, Rice PA. The Modern Epidemic of Syphilis. New England Journal of Medicine 2020;382:845-854. https://doi.org/10.1056/NEJMRA1901593;WGROUP:STRING:MMS.
6. WHO. Global Health Sector Strategies on HIV, Viral Hepatitis and the sexually transmitted infections. World Health Organization 2021;ector stra:2022-2230.
7. Kamb ML, Newman LM, Riley PL, Mark J, Hawkes SJ, Malik T, et al. A Road Map for the Global Elimination of Congenital Syphilis. Obstet Gynecol Int 2010;2010:312798. https://doi.org/10.1155/2010/312798.
8. U.S. Syphilis Cases in Newborns Continue to Increase: A 10-Times Increase Over a Decade | CDC Online Newsroom | CDC n.d. https://www.cdc.gov/media/releases/2023/s1107-newborn-syphilis.html (accessed April 2, 2026).
9. CDC Reports Latest National Data on Syphilis in Newborns and Sexually Transmitted Infections (STIs) | CDC Newsroom n.d. https://www.cdc.gov/media/releases/2025/2025-cdc-reports-latest-national-data-on-syphilis-in-newborns-and-sexually-transmitted-infections-stis.html (accessed April 2, 2026).
10. Moschese D, Lazzarin S, Colombo ML, Caruso F, Giacomelli A, Antinori S, et al. Breakthrough Acute HIV Infections among Pre-Exposure Prophylaxis Users with High Adherence: A Narrative Review. Viruses 2024, Vol 16, 2024;16. https://doi.org/10.3390/V16060951.
11. Seed CR, Styles CE, Hoad VC, Yang H, Thomas MJ, Gosbell IB. Effect of HIV pre-exposure prophylaxis (PrEP) on detection of early infection and its impact on the appropriate post-PrEP deferral period. Vox Sang 2021;116:379-387. https://doi.org/10.1111/VOX.13011.
12. Rodgers MA, Vallari AS, Yamaguchi J, Holzmayer V, Harris B, Toure-Kane C, et al. ARCHITECT HIV Combo Ag/Ab and RealTime HIV-1 Assays Detect Diverse HIV Strains in Clinical Specimens. AIDS Res Hum Retroviruses 2018;34:314. https://doi.org/10.1089/AID.2017.0244.
13. Miedouge M, Grèze M, Bailly A, Izopet J. Analytical sensitivity of four HIV combined antigen/antibody assays using the p24 WHO standard. Journal of Clinical Virology 2011;50:57–60. https://doi.org/10.1016/j.jcv.2010.09.003.
14. Association WM. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA 2013;310:2191-2194. https://doi.org/10.1001/JAMA.2013.281053.

15. International Organization for Standardization (ISO). ISO 15189:2022. Medical laboratories – requirements for quality and competence 2022.
16. Abbott Laboratories Abbott Ireland Diagnostics Division. Alinity i HIV Ag/Ab Combo Reagent Kit, Ref. No. 08P07, Instructions for Use [Package Insert]
17. Mindray Bio-Medical Electronics Co. Ltd. CL-series HIV Antigen and Antibodies to Human Immunodeficiency Virus (CLIA), Ref. No. HIV121, Instructions for Use [Package Insert]
18. Shenzhen New Industries Biomedical Engineering Co. Ltd. MAGLUMI® HIV Ab/Ag Combi (CLIA), Ref. No. 130219008M, Instructions for Use [Package Insert]
19. NIBSC (National Institute for Biological Standards and Control), WHO International Laboratory for Biological Standards, UK Official Medicines Control Laboratory. International Reference Reagent, HIV-1 P24 Antigen, NIBS code: 90/636, Instructions for Use [Package Insert Version 6.0, Dated 12/12/2023].
20. Miedouge M, Grèze M, Bailly A, Izopet J. Analytical sensitivity of four HIV combined antigen/antibody assays using the p24 WHO standard. Journal of Clinical Virology 2011;50:57–60. https://doi.org/10.1016/j.jcv.2010.09.003.
21. Abbott Laboratories Abbott Ireland Diagnostics Division. Alinity i Syphilis Reagent Kit, Ref. No. 07P60, Instructions for Use [Package Insert]
22. Shenzhen New Industries Biomedical Engineering Co. Ltd. MAGLUMI® Syphilis (CLIA), Ref. No. 130269003M, Instructions for Use [Package Insert]
23. Mindray Bio-Medical Electronics Co. Ltd. CL-series Anti-TP, Antibody to Treponema Pallidum (CLIA), Ref. No. 105-005671-00, Instructions for Use [Package Insert]
24. Clinical and Laboratory Standards Institute. CLSI EP15-A3. User Verification of Precision and Estimation of Bias; Approved Guideline - 3rd Edition 2014.
25. NIBSC (National Institute for Biological Standards and Control), WHO International Laboratory for Biological Standards, UK Official Medicines Control Laboratory. WHO Reference Panel, 1st WHO International Reference Panel for HIV-1 p24 Antigen, NIBS code: 16/210, Instructions for Use [Package Insert Version 3.0, Dated 23/09/2024]
26. World Health Organization. WHO Expert Committee on Biological Standardization - Sixty-ninth report, Section 7.1.2 First WHO International Reference Panel for HIV-1 p24 antigen and First WHO International Reference Reagent for HIV-2 p26 antigen. WHO Technical Report Series, No 1016, 2019
27. LLC SD. ASSURANCETM Interference Test Kit INT-06 Instructions for Use [Package Insert]; Sample Interference Testing procedure (based on CLSI EP-7A guidelines), dilution calculator and results reporting spreadsheet
28. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuSetTM HIV-1 Worldwide Performance Panel, Ref. 0800-0492, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuSet-HIV-Worldwide-Performance-Panel-0800-0492/ 2021.
29. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0243, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000243-0600-0243/ 2023.
30. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0245, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000245-0600-0245/ 2023.
31. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0248, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000248-0600-0248/ 2022.
32. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0252, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000252-0600-0252/ 2022.
33. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0258, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000258-0600-0258/ 2024.
34. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0260, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000260-0600-0260/ 2023.
35. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0262, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000262-0600-0262/ 2023.
36. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0265, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000265-0600-0265/ 2021.
37. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0271, Instructions for Use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000271-0600-0271/ 2024.
38. SeraCare LGC Group (LGC Limited. Guildford. UK). AccuVertTM HIV-1 Seroconversion Panel, Ref. 0600-0272, Instructions for use [Package Insert, Data Sheets] https://www.seracare.com/AccuVert-HIV1-Seroconversion-Panel-06000272-0600-0272/ 2022.
39. Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher

RT, Peddada L, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS 2003;17:1871-1879. https://doi.org/10.1097/00002030-200309050-00005.
40. Pandori MW, Hackett J, Louie B, Vallari A, Dowling T, Liska S, et al. Assessment of the Ability of a Fourth-Generation Immunoassay for Human Immunodeficiency Virus (HIV) Antibody and p24 Antigen To Detect both Acute and Recent HIV Infections in a High-Risk Setting. J Clin Microbiol 2009;47:2639. https://doi.org/10.1128/JCM.00119-09.
41. Chéret A. Acute HIV-1 Infection: Paradigm and Singularity. Viruses 2025;17:366. https://doi.org/10.3390/v17030366.
42. Gaspari V, Filippini A, Orioni G, Mussi M, Carpanese MA, La Placa M, et al. The Rise of Syphilis Infections and Reinfections over a Decade (2009–2019) in the Bolognese Area: A Retrospective Analysis. Microorganisms 2025;13:285. https://doi.org/10.3390/microorganisms13020285.
43. Nair M, Gettins L, Fuller M, Kirtley S, Hemelaar J. Global and regional genetic diversity of HIV-1 in 2010–21: systematic review and analysis of prevalence. Lancet Microbe 2024;5. https://doi.org/10.1016/S2666-5247(24)00151-4.
44. Regulation - 2017/746 - EN - Medical Device Regulation - EUR-Lex n.d. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32017R0746 (accessed April 2, 2026).
45. World Health Organization (Geneva). World Health Organization (2024). Consolidated guidelines on differentiated HIV testing services n.d.
46. Totten YR, Hardy BM, Bennett B, Rowlinson MC, Crowe S. Comparative Performance of the Reverse Algorithm Using Architect Syphilis TP Versus the Traditional Algorithm Using Rapid Plasma Reagin in Florida's Public Health Testing Population. Ann Lab Med 2019;39:396-399. https://doi.org/10.3343/ALM.2019.39.4.396.
47. Morshed MG, Singhb AE. Recent trends in the serologic diagnosis of syphilis. Clin Vaccine Immunol 2015;22:137-143. https://doi.org/10.1128/CVI.00681-14.

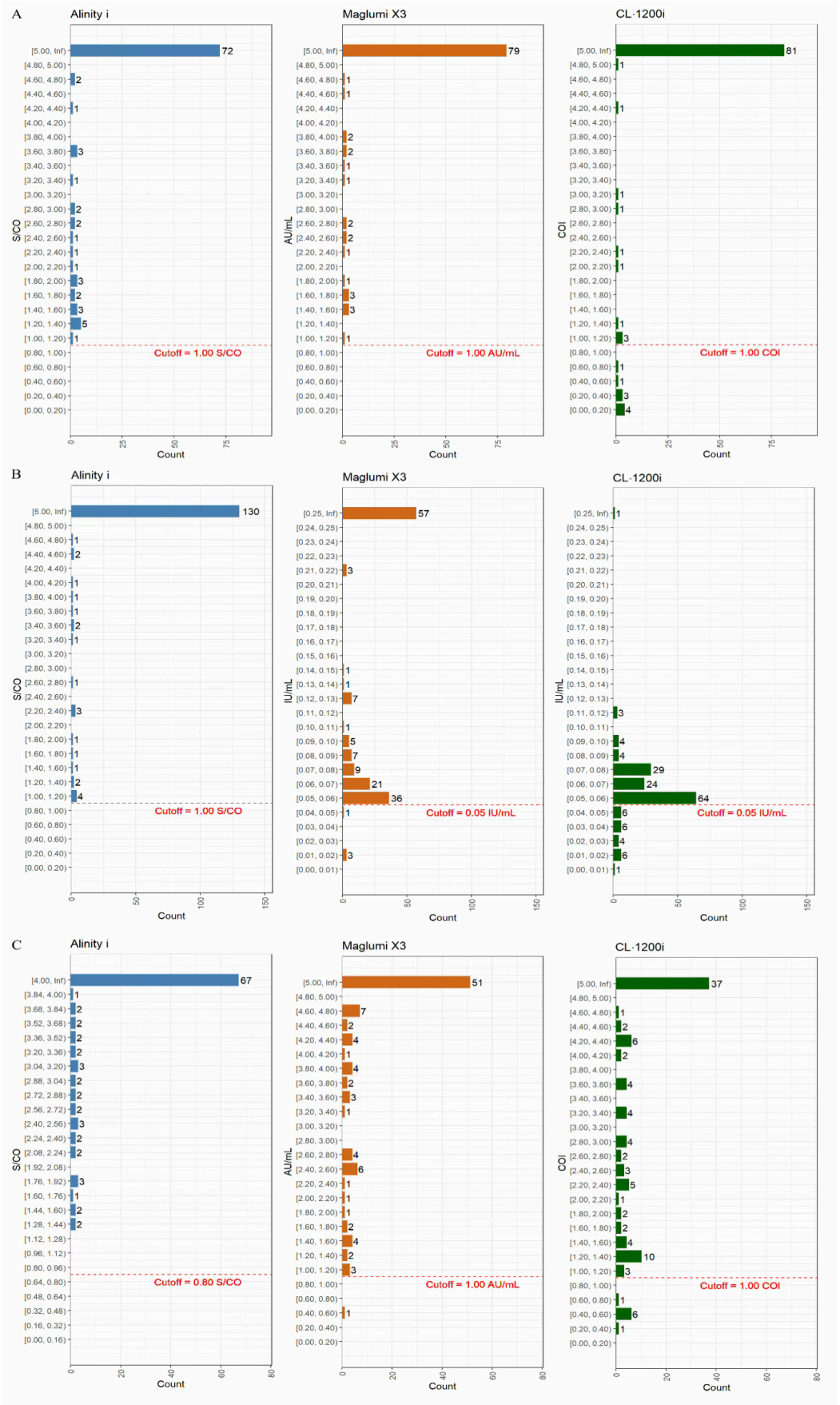
Supplementary Table 1: Characteristics and technical specifications of each system.

	Alinity i	Maglumi X3	CL1200i
Manufacturer	Abbott Laboratories Inc.	Shenzhen New Industries Biomedical Engineering (Snibe) Co., Ltd.	Mindray BioMedical Electronics Co., Ltd.
Analyzer model	Alinity i	Maglumi X3	CL1200i
Method	Chemiluminescent microparticle immunoassay (CMIA)		
Measurement principle	Flash CLIA	Flash CLIA	Glow CLIA
CLIA Label	Acridinium	ABEI (Isoluminol derivate)	ALP
CLIA Substrate	NaOH, H2O2	NaOH, H2O2	AMPPD
Solid phase	paramagnetic microparticles	paramagnetic nanoparticles	paramagnetic microparticles
Capacity	200 tests/hour	200 tests/hour	180 tests/hour

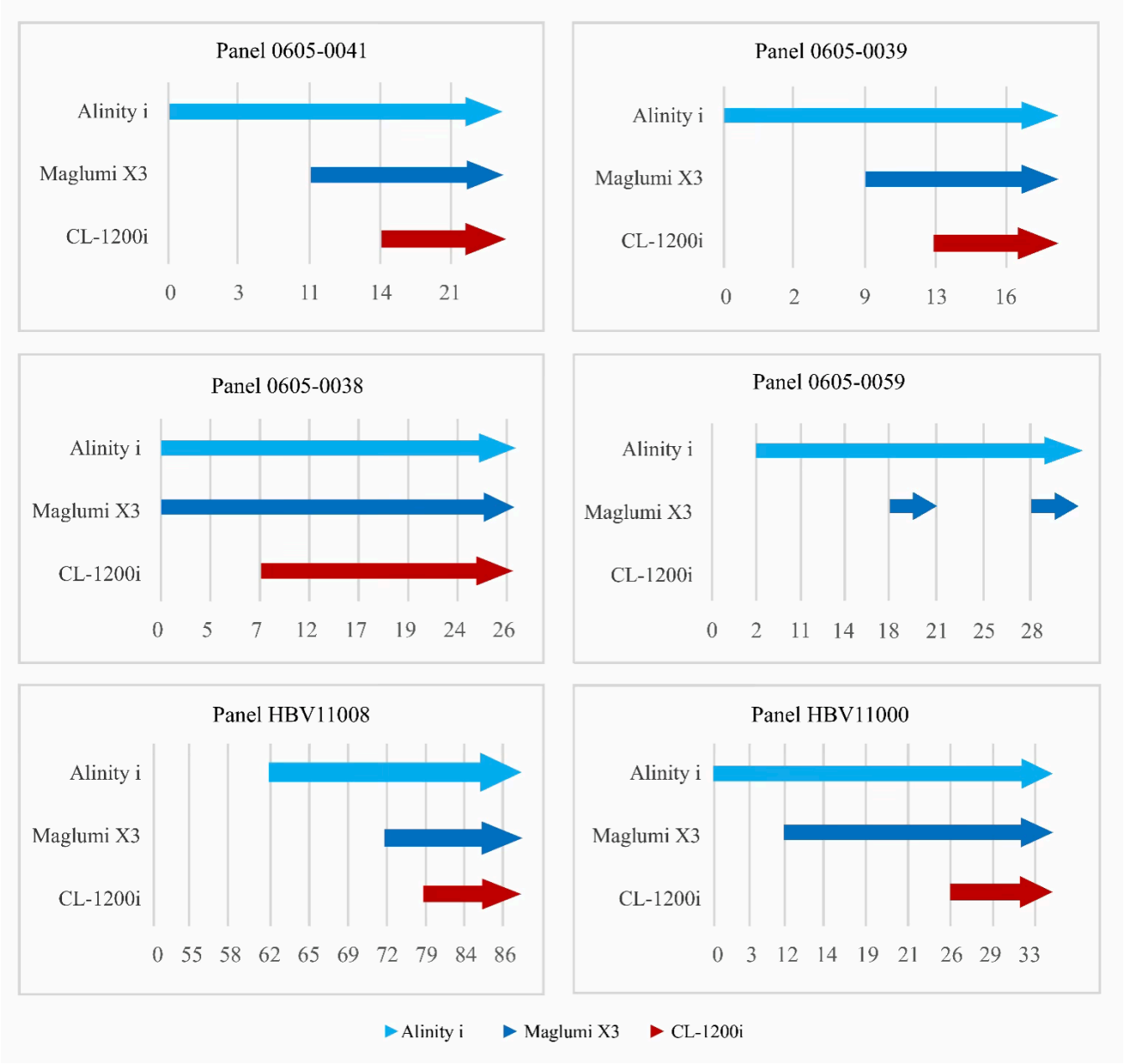
Supplementary Table 2: Precision evaluation results for Anti-HCV, HBsAg and HAV IgM.

Assay	Material	Analyzer	N	Mean	Unit	Repeatability		Within-Lab		95% Lower CL Within-Lab		95% Upper CL Within-Lab	
						(SD)	(CV%)	(SD)	(CV%)	(SD)	(CV%)	(SD)	(CV%)
Anti-HCV	Negative Control (-)	Alinity i	25	0.08	S/CO	0.004	5.35	0.006	7.06	0.004	5.06	0.010	11.66
		Maglumi X3	25	0.31	AU/mL	0.016	5.12	0.165	53.39	0.099	32.07	0.469	151.7
		CL-1200i	25	0.13	COI	0.012	9.38	0.013	9.92	0.010	7.66	0.018	14.07
	Positive Control (+)	Alinity i	25	4.16	S/CO	0.062	1.48	0.066	1.59	0.051	1.22	0.094	2.27
		Maglumi X3	25	9.82	AU/mL	0.382	3.89	0.788	8.02	0.508	5.17	1.736	17.67
		CL-1200i	25	3.28	COI	0.158	4.82	0.158	4.82	0.123	3.76	0.22	6.71
HBsAg	Negative Control (-)	Alinity i	25	0.32	S/CO	0.023	7.08	0.027	8.39	0.020	6.25	0.041	12.77
		Maglumi X3	25	0.03	IU/mL	0.004	14.69	0.005	16.19	0.004	12.37	0.007	23.47
		CL-1200i	25	0.02	IU/mL	0.002	10.70	0.005	28.18	0.003	17.65	0.011	68.31
	Positive Control (+)	Alinity i	25	3.26	S/CO	0.068	2.07	0.068	2.07	0.053	1.61	0.094	2.88
		Maglumi X3	25	1.02	IU/mL	0.036	3.54	0.036	3.58	0.028	2.79	0.051	4.99
		CL-1200i	25	0.57	IU/mL	0.042	7.49	0.042	7.49	0.033	5.84	0.059	10.43
	Positive Control (2)	Maglumi X3	25	100.9	IU/mL	3.258	3.23	4.437	4.40	3.147	3.12	7.512	7.45
HAV IgM	Negative Control (-)	Alinity i	25	0.14	S/CO	0.009	6.80	0.011	8.08	0.008	6.01	0.017	12.34
		Maglumi X3	25	0.54	AU/mL	0.044	8.08	0.045	8.29	0.035	6.44	0.063	11.61
		CL-1200i	25	0.48	COI	0.020	4.17	0.021	4.33	0.016	3.36	0.029	6.09
	Positive Control (+)	Alinity i	25	1.81	S/CO	0.061	3.39	0.062	3.43	0.048	2.67	0.087	4.78
		Maglumi X3	25	4.4	AU/mL	0.172	3.90	0.172	3.90	0.134	3.04	0.239	5.43
		CL-1200i	25	2.92	COI	0.133	4.55	0.133	4.55	0.103	3.55	0.185	6.33

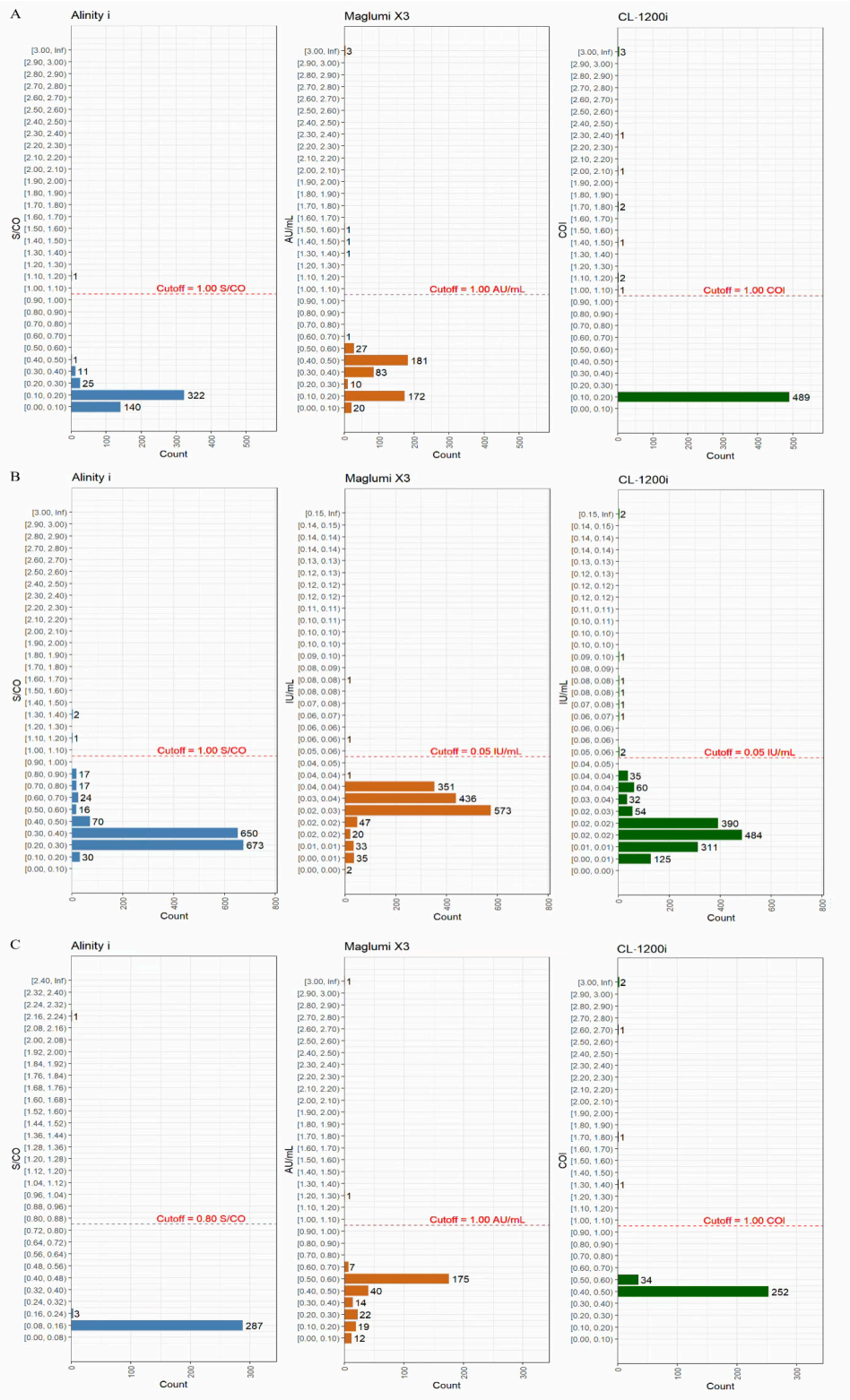
Supplementary Figure 1: Clinical sensitivity evaluation results distribution on the three CLIA platforms. (A) Anti-HCV; (B) HBsAg (C) HAV IgM.



Supplementary Figure 2: Number of days since the 1st HBsAg positive test result detection. Colour coded arrows are shown for each of the analyzers (panel members tested negative between first and last positive detection, are presented by separate arrows, indicating the intermittent detection). Arrows heads exceed the last detection days to reflect that the sample will be detected as positive from then onwards.



Supplementary Figure 3: Clinical specificity evaluation results distribution on the three CLIA platforms. (A) Anti-HCV; (B) HBsAg (C) HAV IgM.



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

Comparison of Cardiovascular Risk Classification by Established Risk Scores Across Glycemic Status in an Indian Population

Lathangi G[1], Pooja R[1], Nisarga T[1], Manish Kumar Pandey[1], Jadeppa G[1], Swetha N. Kempegowda[1*]

[1]*Department of Biochemistry, Centre of Excellence in Molecular Biology, JSS Medical College and Hospital, JSS Academy of Higher Education & Research, Mysore, Karnataka, India*

Article Info

[*]***Corresponding Author:***
Swetha N.K
Department of Biochemistry
JSS Medical College
JSS Academy of Higher Education and Research
Mysore
E-mail address: swethank@jssuni.edu.in
ORCID ID:0000-0002-3163-7405

Keywords

Cardiovascular risk, Framingham Risk Score, ASCVD Risk Estimator Plus, QRISK3, WHO non-laboratory risk charts, Clinical Utility Index, Classification concordance, India

Abstract

Background: Cardiovascular disease (CVD) risk estimation is central to preventive care, particularly in populations with a high burden of diabetes and prediabetes. Most widely used cardiovascular risk prediction models were developed in Western populations and may show variable risk classification when applied to Indian cohorts.

Objectives: To compare cardiovascular risk classification using four established 10-year risk prediction models - Framingham Risk Score (FRS), ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory risk charts - across different glycemic statuses in an Indian population.

Methods: This cross-sectional analytical study included 257 adults classified as normoglycemic controls (n = 76), prediabetics (n = 49), and diabetics (n = 132). Ten-year cardiovascular risk was estimated using FRS, ASCVD, QRISK3, and WHO non-laboratory charts. Risk category distributions were compared using chi-square analysis. Comparative classification behaviour across models was assessed using ROC curves, comparative classification metrics, and Clinical Utility Index (CUI), interpreted solely as measures of comparative classification behaviour rather than predictive validation.

Results: Substantial heterogeneity in cardiovascular risk classification was observed across models. QRISK3 classified the highest proportion of diabetic individuals as high cardiovascular risk (40.9%), whereas FRS classified only 6.8% into the high-risk category. WHO non-laboratory charts did not identify any participant as high risk at the ≥30% threshold. Significant differences across glycemic groups were observed for FRS, ASCVD, and QRISK3 ($p < 0.05$). QRISK3 demonstrated a greater tendency to classify participants into higher cardiovascular risk categories, whereas FRS and WHO non-laboratory charts showed a more conservative classification pattern.

Conclusion: Cardiovascular risk classification varies substantially across commonly used prediction models when applied to Indian individuals across different glycemic states. These findings reflect differences in classification behaviour rather than predictive performance and highlight the need for population-specific validation using longitudinal cardiovascular outcome studies.

Key Message

- Existing cardiovascular risk models developed in Western populations may not accurately reflect risk in Indian populations.
- This study found marked variability in risk classification across FRS, ASCVD, QRISK3, and WHO models, particularly among diabetics.
- The findings highlight the need for cautious interpretation and population-specific validation of cardiovascular risk assessment tools in India.

Introduction

Cardiovascular disease (CVD) remains the leading cause of global mortality, accounting for nearly one-third of all deaths worldwide, with a disproportionate burden in low- and middle-income countries such as India [1]. Diabetes mellitus accelerates atherosclerosis through endothelial dysfunction, chronic inflammation, oxidative stress, dyslipidemia, and prothrombotic mechanisms, substantially increasing cardiovascular morbidity and mortality [2]. Prediabetes is now recognized as an independent cardiovascular risk state, with evidence of early vascular dysfunction and subclinical atherosclerosis preceding overt diabetes [3].

Indian populations exhibit a distinctive cardiometabolic phenotype characterized by higher visceral adiposity, insulin resistance, low high-density lipoprotein cholesterol, elevated triglycerides, and earlier onset of cardiovascular events at lower body mass indices compared with Western populations [4]. This "South Asian phenotype" contributes to premature coronary artery disease and higher case-fatality rates, underscoring the need for accurate cardiovascular risk stratification in Indian settings [5]. Despite this vulnerability, cardiovascular risk assessment in India largely relies on prediction models developed in Western cohorts, which may inadequately capture ethnicity-specific risk patterns [6].

Estimation of 10-year cardiovascular risk is central to preventive cardiology and informs decisions regarding lifestyle modification and pharmacological therapy [7]. The Framingham Risk Score (FRS) and ASCVD Risk Estimator Plus are widely used but are primarily validated in Western populations and may underestimate cardiovascular risk in South Asians [8,9].QRISK3 incorporates a broader range of variables, including ethnicity and comorbid conditions, potentially improving risk classification in diverse populations, although it was developed using United Kingdom data and remains inadequately evaluated in resident Indian populations [10,11]. The World Health Organization (WHO) non-laboratory cardiovascular risk charts facilitate risk estimation in low-resource settings but may underestimate risk among high-risk individuals [12].

Previous Indian studies have demonstrated substantial discordance among cardiovascular risk prediction models, with some tools classifying individuals with established cardiovascular disease as low risk, raising concerns regarding their clinical applicability [13,14].

However, there is limited comparative evidence evaluating cardiovascular risk classification and clinical utility of multiple risk prediction models across different glycemic states - diabetes, prediabetes, and normoglycemia - in Indian populations [15]. Clinical Utility Index (CUI) and receiver operating characteristic (ROC) curve analysis provide complementary approaches for evaluating the practical and discriminatory characteristics of risk prediction models beyond conventional statistical measures [16–18].

Therefore, the present study was undertaken to compare cardiovascular risk classification using FRS, ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory risk charts among diabetic, prediabetic, and normoglycemic individuals in an Indian population, with the aim of informing appropriate cardiovascular risk assessment in Indian clinical practice.

Methods

Study Setting and Design

This was a cross-sectional analytical study conducted at a tertiary-care teaching hospital in South India. The hospital serves a heterogeneous population from both urban and rural backgrounds, making it suitable for evaluating cardiovascular risk stratification across different glycemic states. The study was designed to compare cardiovascular risk classification patterns generated by multiple established 10-year cardiovascular risk prediction models within the same cohort.

Sample Size

The sample size was selected to support a comparative evaluation of cardiovascular risk classification across multiple prediction models in a cross-sectional design. A total of 257 participants is comparable to prior Indian studies assessing inter-model variability in cardiovascular risk scores and is adequate for detecting differences in risk category distribution using chi-square analysis and estimating relative discriminatory behaviour through ROC analysis. No formal a priori sample size calculation was performed because this exploratory cross-sectional study included all eligible participants meeting the predefined inclusion criteria during the study period. Therefore, the available sample of 257 participants was considered appropriate for hypothesis-generating comparison of cardiovascular risk classification patterns rather than formal model validation [6,11].

Study Population and Sampling Strategy

Adult participants attending outpatient departments and inpatient services of the hospital during the study period were screened for eligibility. A consecutive sampling strategy was employed to recruit eligible participants who met the inclusion criteria. Individuals were enrolled after obtaining informed written consent. Participants were categorized into three groups based on glycemic status: healthy controls, prediabetics, and diabetics.

Inclusion Criteria

1. Adults aged ≥18 years
2. Willingness to provide informed consent
3. Availability of clinical and biochemical data required for cardiovascular risk calculation

Exclusion Criteria

Known history of established cardiovascular disease (e.g., myocardial infarction, stroke, coronary revascularization)

Operational Definitions

Diabetes mellitus was defined according to the American Diabetes Association (ADA) criteria, based on fasting plasma glucose, HbA1c values, or documented medical history. Prediabetes was defined as impaired fasting glucose and/or HbA1c values within the prediabetic range as per ADA guidelines. Cardiovascular risk categorization was performed using four established risk prediction models. High cardiovascular risk was defined as a 10-year risk ≥20% for Framingham Risk Score (FRS), ASCVD Risk Estimator Plus, and QRISK3, while a threshold of ≥30% was used for the WHO non-laboratory risk charts, in accordance with standard recommendations for each model.

Study Procedures

Detailed demographic and clinical information was collected from all participants using a structured proforma. Data recorded included age, sex, smoking status, family history of cardiovascular disease, and relevant medical history. Medication history including antihypertensive therapy, lipid-lowering agents, glucose-lowering medications, and antiplatelet therapy was recorded whenever available and incorporated into the corresponding cardiovascular risk score calculations according to the requirements of each model. Anthropometric measurements such as height and weight were obtained using calibrated instruments, and body mass index (BMI) was calculated. Blood pressure was measured using a standardized protocol after adequate rest. Biochemical parameters, including fasting plasma glucose and lipid profile (total cholesterol, HDL cholesterol, LDL cholesterol, and triglycerides) were obtained from laboratory records of the Department of Biochemistry at the study institution. All biochemical analyses were performed in the same laboratory using standardized analytical procedures and calibrated automated analyzers according to institutional quality control protocols. Fasting plasma glucose and lipid parameters were measured using routine enzymatic methods employed in the clinical laboratory.

The baseline demographic, anthropometric, clinical, and biochemical characteristics of the study participants according to glycemic status are summarized in Table 1.

Table 1: Baseline demographic, anthropometric, clinical, and biochemical characteristics of study participants according to glycemic status.

Variable	Control (n=76)	Prediabetes (n=49)	Diabetes (n=132)	p-value
Age (years)	38.48 ± 11.72	52.33 ± 13.47	55.90 ± 12.39	<0.0001
Males n (%)	32 (42.1)	24 (49.0)	63 (47.7)	0.86
Female n (%)	44 (57.9)	25 (51.0)	69 (52.3)	
Current smokers n (%)	4 (5.3)	7 (14.3)	15 (11.4)	0.21
Non-smokers n (%)	72 (94.7)	42 (85.7)	117 (88.6)	
Height (cm)	164.1 ± 10.13	159.7 ± 11.04	161.8 ± 10.23	0.0557
Weight (kg)	68.28 ± 11.83	70.65 ± 13.42	70.42 ± 13.33	0.4247
BMI (kg/m²)	25.36 ± 3.74	27.87 ± 4.70	26.94 ± 4.90	0.0039
Waist-Hip Ratio	1.055 ± 0.22	0.9525 ± 0.08	0.9455 ± 0.10	<0.0001
Systolic Blood Pressure (mmHg)	119.8 ± 16.87	133.4 ± 18.13	133.1 ± 19.90	<0.0001
Diastolic Blood Pressure (mmHg)	79.49 ± 10.99	85.0 ± 10.27	81.91 ± 10.78	0.0161
Fasting Plasma Glucose (mg/dL)	90.09 ± 24.57	98.0 ± 13.68	161.7 ± 69.33	<0.0001
HbA1c (%)	5.393 ± 0.66	5.949 ± 0.20	8.49 ± 1.76	<0.0001

Variable	Control (n=76)	Prediabetes (n=49)	Diabetes (n=132)	p-value
Total Cholesterol (mg/dL)	169.4 ± 36.26	176.4 ± 42.06	172.7 ± 41.35	0.603
Triglycerides (mg/dL)	124.5 ± 67.56	145.7 ± 66.27	170.3 ± 101.70	0.0007
HDL Cholesterol (mg/dL)	42.21 ± 8.79	42.82 ± 17.08	42.72 ± 11.75	0.9439
LDL Cholesterol (mg/dL)	115.9 ± 31.87	119.9 ± 39.58	116.2 ± 40.41	0.8144

Data are presented as mean ± standard deviation (SD) for continuous variables and frequency (percentage) for categorical variables. P-values were obtained using one-way ANOVA for continuous variables and chi-square test for categorical variables.

Significant differences were observed across glycemic groups for age, body mass index, blood pressure, fasting plasma glucose, HbA1c, and triglyceride levels, whereas sex distribution, smoking status, and most lipid parameters were comparable between groups. These baseline characteristics provide important context for interpreting differences in cardiovascular risk classification across the evaluated risk prediction models.

Cardiovascular Risk Prediction Models

For each participant, 10-year cardiovascular risk was independently calculated using the Framingham Risk Score, ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory risk charts, following the input variables and cut-off thresholds specified for each model.The Framingham Risk Score (FRS), ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory cardiovascular risk charts differ in the clinical and metabolic variables incorporated for cardiovascular risk estimation. All four models include age, sex, systolic blood pressure, smoking status, and diabetes status. Ethnicity is incorporated in ASCVD and QRISK3, while body mass index is included only in QRISK3 and WHO non-laboratory models. Lipid parameters such as total cholesterol and high-density lipoprotein cholesterol are used in FRS and ASCVD, whereas QRISK3 employs the cholesterol-to-HDL ratio and WHO non-laboratory charts do not include biochemical lipid variables. QRISK3 uniquely incorporates additional clinical comorbidities and treatment-related variables, including family history of premature cardiovascular disease, chronic kidney disease, atrial fibrillation, inflammatory disorders, medication use, and socioeconomic deprivation indices (Table 2). Variables unique to QRISK3 including ethnicity, family history, chronic kidney disease, medication use, and comorbid conditions were operationalized using available clinical records. Variables unavailable in the study cohort were handled according to model guidance and documented assumptions.

Table 2: Comparative distribution of clinical and metabolic parameters included in major cardiovascular risk prediction models (FRS, ASCVD, QRISK3, and WHO non-laboratory charts).

Parameter	FRS	ASCVD	QRISK3	WHO Non-Lab
Age	✓	✓	✓	✓
Sex	✓	✓	✓	✓
Ethnicity / Race	X	✓ (race)	✓	X
Total Cholesterol	✓	✓	X	X
HDL Cholesterol	✓	✓	X (uses ratio)	X
Chol/HDL Ratio	X	X	✓	X
LDL Cholesterol	X	✓	X	X
Systolic Blood Pressure	✓	✓	✓	✓
Hypertension Treatment	✓	✓	X	X
Smoking Status	✓	✓	✓	✓
Diabetes Status	✓	✓	✓ (Type 1 / 2)	✓
BMI	X	X	✓	✓

Parameter	FRS	ASCVD	QRISK3	WHO Non-Lab
Family History of CVD	X	X	✓ (<60 yrs)	X
Chronic Kidney Disease	X	X	✓	X
Atrial Fibrillation	X	X	✓	X
Rheumatoid Arthritis	X	X	✓	X
Migraine	X	X	✓	X
Severe Mental Illness	X	X	✓	X
Systemic Lupus Erythematosus	X	X	✓	X
Corticosteroid Use	X	X	✓	X
Antipsychotic Use	X	X	✓	X
Socioeconomic / Deprivation Index	X	X	✓	X

✓- *Clinical and Metabolic Parameters Included Under Each Risk Score*

X- Clinical and Metabolic Parameters Not Included Under Each Risk Score

Data Management and Statistical Analysis

Data were entered into a structured database and analyzed using SPSS and GraphPad Prism software. Categorical variables were summarized as frequencies and percentages. Differences in cardiovascular risk category distribution across glycemic groups were assessed using the chi-square test.

To compare the relative classification behaviour of the risk models within the cohort, Relative High-Risk Classification Concordance, Relative Low-Risk Classification Concordance, Positive Classification Concordance, Negative Classification Concordance, the area under the receiver operating characteristic (ROC) curve (AUC), and the Clinical Utility Index (CUI) were calculated using binary high-risk versus non-high-risk classification. For comparative classification analyses, diabetes mellitus status was used as the reference classification state because longitudinal cardiovascular outcome data were unavailable. Diabetes mellitus was selected as the surrogate reference classification because it represents one of the strongest established determinants of future cardiovascular disease and is incorporated as a major predictor in all evaluated cardiovascular risk models. In the absence of longitudinal cardiovascular outcome data, diabetes status provided a clinically relevant comparator for evaluating relative differences in cardiovascular risk classification across the models rather than true prediction of cardiovascular events. Accordingly, ROC analysis was used solely to compare relative classification behaviour across models and should not be interpreted as validation against future cardiovascular events. These measures were interpreted as comparative classification metrics, rather than as validation against clinical cardiovascular outcomes, as no longitudinal event data were available.

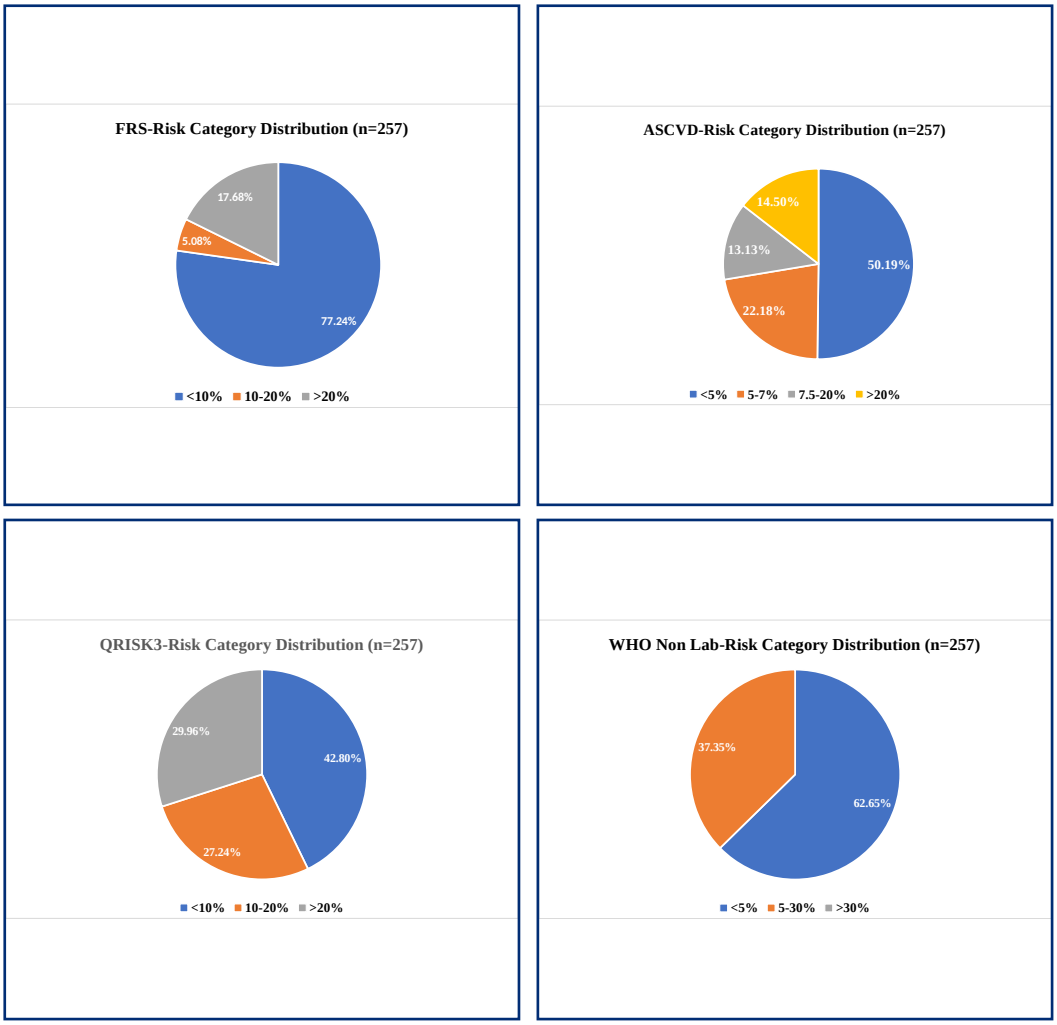
Clinical Utility Index for positive (CUI^+) and negative (CUI^-) outcomes was calculated to assess the practical implications of risk classification in identifying individuals more likely to be classified as high or low risk by each model. A p-value <0.05 was considered statistically significant for all analyses. As longitudinal cardiovascular outcomes were unavailable, Relative High-Risk Classification Concordance, Relative Low-Risk Classification Concordance, Positive Classification Concordance, Negative Classification Concordance, ROC analysis, and Clinical Utility Index were used only as comparative measures describing differences in classification behaviour across models and should not be interpreted as measures of predictive accuracy. Participants with incomplete data required for risk score calculation were excluded from model-specific analyses. No imputation procedure was applied.

Results

Risk Category Distribution Across Cardiovascular Risk Scores

Marked variability was observed in cardiovascular risk classification across the four models (Figure 1). According to the Framingham Risk Score (FRS), only 5.08% of participants were classified as having high cardiovascular risk (>20%), with most participants categorized as low or intermediate risk. The ASCVD Risk Estimator Plus classified 14.5% of participants as high risk (>20%). QRISK3 demonstrated a greater tendency to classify participants into higher risk categories; however, percentages presented in Figure 1 represent distribution across the entire study cohort (n = 257). WHO non-laboratory risk charts did not classify any participant as having a 10-year cardiovascular risk ≥30%.

Figure 1: Distribution of 10-year cardiovascular risk categories across the entire study population (n = 257) using Framingham Risk Score (FRS), ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory risk charts.



Comparison of Cardiovascular Risk Classification Across Glycemic States

Comparison of cardiovascular risk classification across normoglycemic controls, prediabetic individuals, and diabetic individuals demonstrated statistically significant differences for multiple cardiovascular risk prediction models (Table 3). The Framingham Risk Score (FRS) showed a significant variation in risk category distribution across the three glycemic groups ($\chi^2 = 8.852$; $p = 0.012$), with a higher proportion of participants in the prediabetic and diabetic groups classified as high risk compared with controls. Similarly, the ASCVD Risk Estimator Plus demonstrated highly significant differences in cardiovascular risk classification across glycemic states ($\chi^2 = 19.12$; $p < 0.0001$). The QRISK3 model also showed significant variation in cardiovascular risk distribution among controls, prediabetics, and diabetics ($\chi^2 = 28.98$; $p < 0.0001$), with a greater proportion of diabetic individuals categorized into the high-risk group. In contrast, WHO non-laboratory cardiovascular risk charts did not demonstrate variation across glycemic groups, as no participant exceeded the predefined ≥30% cardiovascular risk threshold.

Table 3: Distribution of 10-year cardiovascular risk categories across normoglycemic, prediabetic, and diabetic individuals using FRS, ASCVD, QRISK3, and WHO non-laboratory risk prediction models, with chi-square comparison.

RISK SCORES		CONTROL		PREDIABETIC		DIABETES		CHI SQUARE	p-value
		Frequency	%	Frequency	%	Frequency	%		
FRS	<20	76	100.0	43	87.8	123	93.2	8.852	0.012
	>20	0	0.0	6	12.2	9	6.8		
ASCVD	<20	75	98.7	43	87.8	101	76.5	19.12	<0.0001
	>20	1	1.3	6	12.2	31	23.5		
QRISK3	<20	70	92.1	40	81.6	78	59.1	28.98	<0.0001
	>20	6	7.9	9	18.4	54	40.9		
WHO NONLAB	<30	76	100.0	49	100.0	132	100.0	0	–
	>30	0	0.0	0	0.0	0	0.0		

FRS: Framingham Risk Score; ASCVD: ASCVD Risk Estimator Plus; QRISK3: QRISK3; WHO NONLAB: WHO non-laboratory risk charts. Percentages are calculated within each glycemic status group (column percentages).

Table 4: Comparative Classification Concordance Metrics of FRS, ASCVD, QRISK3, and WHO Non-Laboratory Cardiovascular Risk Prediction Models.

COMPARATIVE CLASSIFICATION MEASURES	FRS >20%		ASCVD >20%		QRISK3 >20%		WHO NONLAB >30%	
	VALUE	95%CI	VALUE	95%CI	VALUE	95%CI	VALUE	95%CI
Relative High-Risk Classification Concordance	6.82	3.16–12.65	23.48	16.62–31.50	40.91	32.37–49.87	0.00	0.00–2.76
Relative Low-Risk Classification Concordance	95.20	88.15–98.69	93.60	85.86–97.44	85.60	76.57–91.68	100.00	94.20–100.0
Positive Classification Concordance	60.00	31.32–83.16	81.58	64.98–92.04	78.26	63.75–88.90	–	–
Negative Classification Concordance	43.07	34.41–51.99	50.43	41.59–59.20	53.73	44.68–62.61	41.25	33.20–49.72

FRS: Framingham Risk Score; ASCVD: ASCVD Risk Estimator Plus; QRISK3: QRISK3; WHO NONLAB: WHO non-laboratory risk charts; CI: Confidence Interval.

Note: Positive clinical state = Diabetes mellitus; Positive test = High cardiovascular risk category as per model thresholds. Diabetes mellitus was used as the surrogate reference classification state because longitudinal cardiovascular outcome data were unavailable. Relative High-Risk Classification Concordance, Relative Low-Risk Classification Concordance, Positive Classification Concordance, and Negative Classification Concordance describe the comparative classification behaviour of the evaluated cardiovascular risk models and should not be interpreted as measures of diagnostic accuracy or predictive performance.

Comparative Classification Behaviour Across Cardiovascular Risk Prediction Models

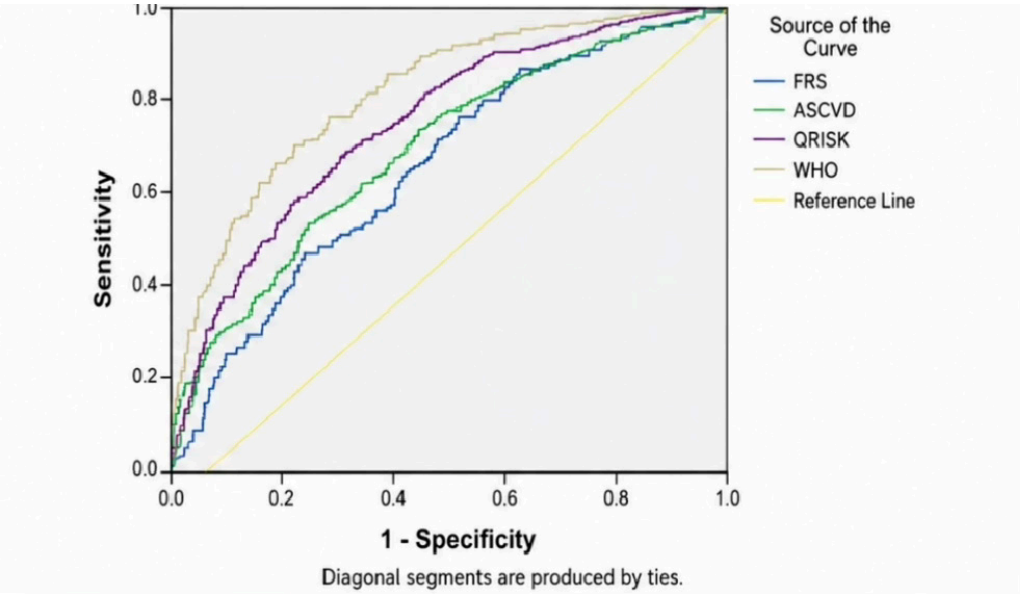
The comparative classification characteristics of the four cardiovascular risk prediction models varied considerably with respect to relative classification concordance for higher- and lower-risk categorization (Table 4). When diabetes mellitus was used as the surrogate reference classification state, QRISK3 demonstrated the greatest concordance for identifying participants classified into higher cardiovascular risk categories, whereas FRS demonstrated greater concordance for classifying participants into lower-risk categories. ASCVD showed intermediate classification concordance, while WHO non-laboratory charts demonstrated no concordance for high-risk classification at the ≥30% threshold because no participant exceeded this cut-off. These findings should be interpreted solely as comparative classification behaviour rather than diagnostic accuracy or predictive performance.

ROC Analysis and Clinical Utility of Cardiovascular Risk Scores

Receiver operating characteristic (ROC) curve analysis was performed to assess the relative discriminatory behaviour of the

cardiovascular risk prediction models in classifying individuals as high versus non-high cardiovascular risk (Figure 2). All four models demonstrated discriminatory ability above the reference line, with QRISK3 showing the highest area under the curve (AUC), followed by ASCVD, WHO non-laboratory charts, and FRS. As the study lacked longitudinal cardiovascular outcome data, ROC analysis was interpreted as a comparative measure of classification behaviour rather than as an assessment of predictive accuracy for incident cardiovascular events.

Figure 2: Receiver operating characteristic (ROC) curves illustrating the comparative classification behaviour of FRS, ASCVD, QRISK3, and WHO non-laboratory cardiovascular risk models for classification of high versus non-high cardiovascular risk, together with corresponding area under the curve (AUC) values and 95% confidence intervals.



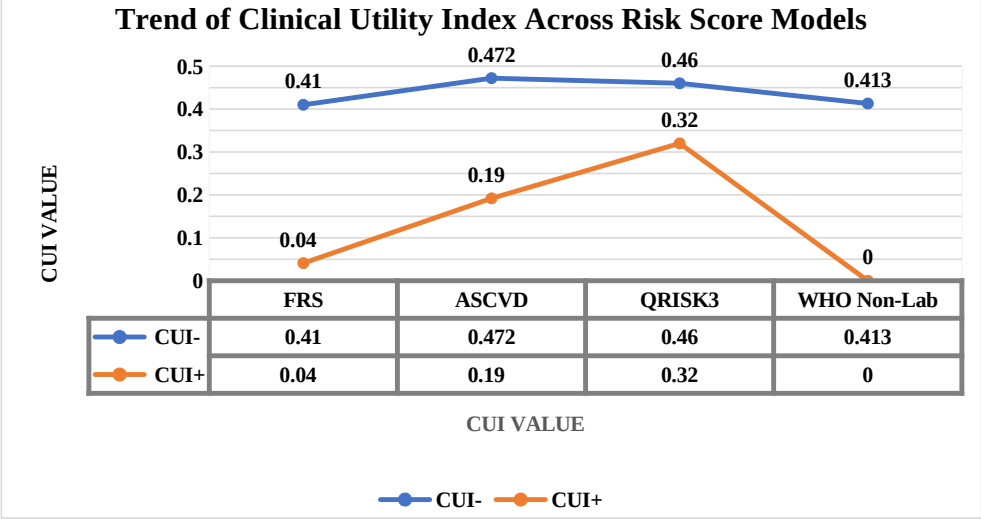
Risk Score	AUC	Standard Error	95% CI		p-value
			Lower Bound	Upper Bound	
FRS	0.733	0.035	0.664	0.801	<0.001
ASCVD	0.729	0.033	0.664	0.794	<0.001
QRISK	0.780	0.031	0.720	0.840	<0.001
WHO	0.787	0.031	0.727	0.847	<0.001

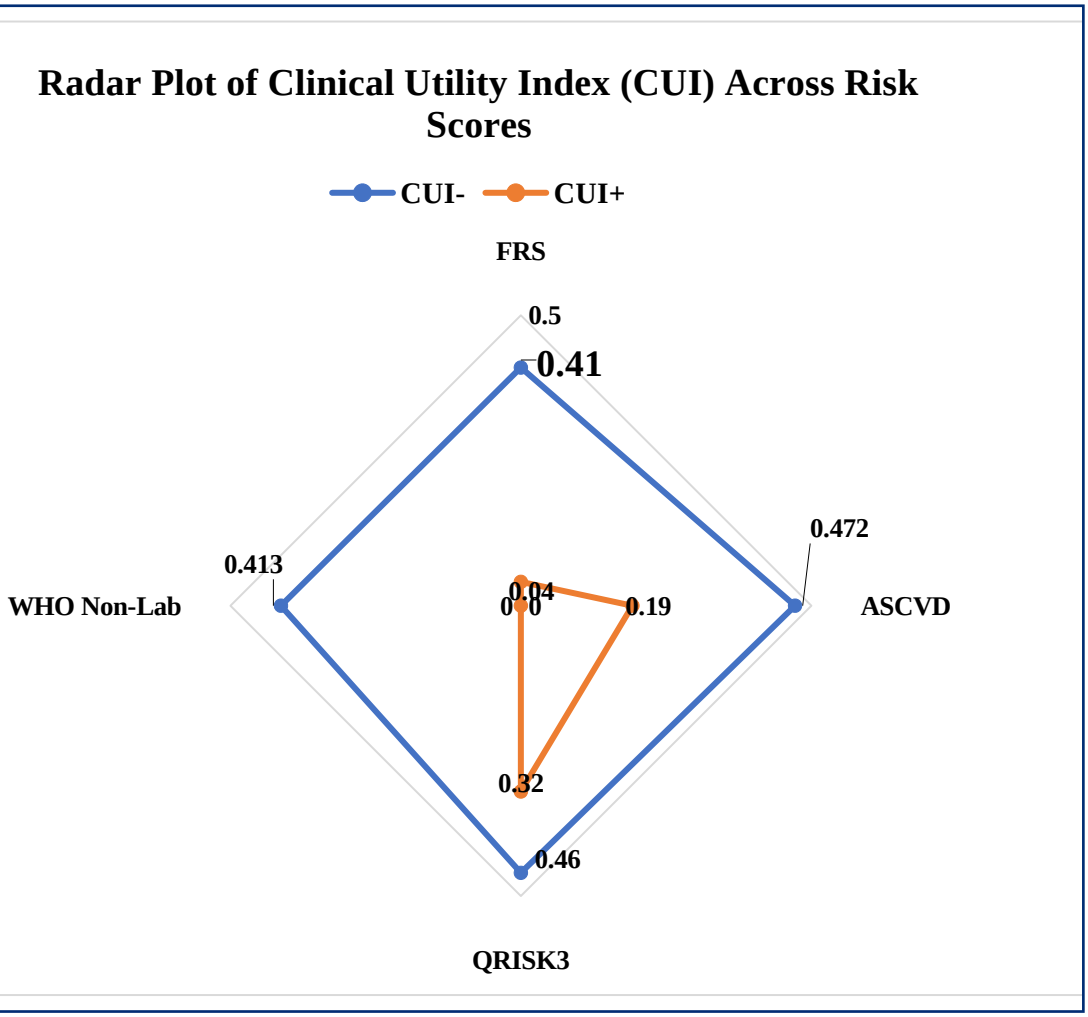
Note: AUC values were estimated using nonparametric ROC analysis.
p-values indicate comparison against the null hypothesis of AUC = 0.5.

Comparison of Clinical Utility Index (CUI^+ and CUI^-) Across Cardiovascular Risk Prediction Models

Clinical Utility Index (CUI) analysis was used to evaluate the practical implications of cardiovascular risk classification by integrating the comparative classification concordance measures. QRISK3 demonstrated the highest positive clinical utility ($CUI^+ = 0.320$), indicating a greater tendency to classify individuals into higher cardiovascular risk categories. In contrast, ASCVD demonstrated the highest negative clinical utility ($CUI^- = 0.472$), indicating a greater tendency to classify participants into lower-risk categories. FRS showed comparatively low positive clinical utility ($CUI^+ = 0.041$) but moderate negative clinical utility ($CUI^- = 0.410$), reflecting a conservative classification profile. WHO non-laboratory cardiovascular risk charts demonstrated no positive clinical utility at the predefined threshold ($CUI^+ = 0.000$), although negative clinical utility remained moderate ($CUI^- = 0.413$). (Figure 3) These findings reflect differences in classification behaviour rather than predictive accuracy for future cardiovascular outcomes.

Figure 3: Trend and radar plot illustrating the comparative clinical utility index for positive (CUI^+) and negative (CUI^-) outcomes across FRS, ASCVD, QRISK3, and WHO non-laboratory cardiovascular risk prediction models.





Discussion

The present study provides a comparative evaluation of four widely used 10-year cardiovascular risk prediction models - Framingham Risk Score (FRS), ASCVD Risk Estimator Plus, QRISK3, and WHO non-laboratory risk charts - across normoglycemic, prediabetic, and diabetic individuals in an Indian population. The principal finding is the substantial heterogeneity in risk classification across models, with marked discordance in the proportion of individuals categorized as high cardiovascular risk, particularly among those with diabetes mellitus. Differences in cardiovascular risk classification may influence how individuals are categorized by commonly used risk assessment tools. In high-burden settings such as India, understanding these classification differences is important when interpreting estimated cardiovascular risk.

Interpretation of Risk Classification Differences

The observed variability across risk scores highlights the challenges of applying cardiovascular risk algorithms developed in Western populations to Indian settings. In the present cohort, QRISK3 classified a considerably higher proportion of diabetic individuals as high risk compared with FRS and ASCVD, while the WHO non-laboratory charts failed to identify any participant as high risk at the recommended ≥30% threshold. These findings underscore that different models emphasize distinct combinations of demographic, metabolic, and clinical variables, leading to divergent risk stratification even within the same population.

QRISK3 incorporates a broader range of variables, including ethnicity, body mass index, family history of premature cardiovascular disease, chronic kidney disease, and selected comorbid conditions. This broader clinical scope may explain its tendency to classify a larger proportion of individuals—particularly diabetics—as higher risk in the present study.

In contrast, FRS and ASCVD rely more heavily on age and traditional lipid parameters, which may lead to relative under-classification of risk in younger individuals with adverse metabolic profiles, a pattern commonly observed in South Asian populations.

Diagnostic Metrics and Methodological Considerations

The classification concordance measures reported in this study (formerly corresponding to sensitivity, specificity, positive predictive value, and negative predictive value), together with the area under the receiver operating characteristic (ROC) curve (AUC) and the Clinical Utility Index (CUI), were used solely for comparative classification analysis within the study cohort and should not be interpreted as measures of diagnostic or prognostic performance. However, it is important to emphasize that this was a cross-sectional study without longitudinal cardiovascular outcomes, and therefore these metrics should not be interpreted as true diagnostic accuracy against a gold-standard clinical endpoint. Rather, they reflect how differently each model classifies individuals as higher or lower risk relative to the others.

Within this context, QRISK3 demonstrated greater relative classification concordance for identifying participants categorized as high cardiovascular risk, whereas FRS and WHO non-laboratory models demonstrated greater concordance for classifying participants into lower-risk categories. These patterns illustrate the trade-off between broader risk capture and conservative classification inherent to different algorithms and should be interpreted as classification tendencies rather than predictive superiority.

WHO Non-Laboratory Risk Charts

A notable finding was the inability of the WHO non-laboratory risk charts to classify any participant as high risk using the recommended ≥30% threshold. This observation raises concerns regarding the suitability of these thresholds in Indian populations, particularly among individuals with diabetes and prediabetes. While WHO charts are designed for low-resource settings and prioritize simplicity, their limited sensitivity in this cohort suggests potential underestimation of cardiovascular risk in high-burden populations. These findings are consistent with previous Indian studies reporting under-classification of cardiovascular risk by WHO risk charts and highlight the need for cautious interpretation during clinical decision-making in high-risk populations [11,14].

The absence of participants classified at the ≥30% threshold may also reflect the relatively conservative threshold recommended for WHO non-laboratory risk charts rather than complete failure of the model itself. Alternative thresholds such as ≥20% were not explored because the objective of the present study was to compare each model using its recommended high-risk threshold. Future studies should evaluate whether lower WHO thresholds improve cardiovascular risk classification in Indian populations.

Clinical Utility Perspective

The incorporation of the Clinical Utility Index (CUI) allowed assessment of the practical implications of cardiovascular risk classification beyond conventional statistical measures. QRISK3 demonstrated higher positive clinical utility, indicating a greater tendency to classify participants into higher cardiovascular risk categories. Conversely, FRS and WHO non-laboratory models demonstrated a greater tendency to classify participants into lower-risk categories. These findings should be interpreted as differences in classification behaviour rather than evidence of predictive superiority or clinical effectiveness. This complementary interpretation reinforces the concept that no single risk score is universally optimal and that the choice of risk assessment tool may influence cardiovascular risk categorization within the same population.

The findings of this study are consistent with previous Indian and South Asian studies demonstrating discordance among cardiovascular risk prediction models and underestimation of risk by traditional Western-derived tools [6]. Prior

investigations in Indian cohorts, including patients with established coronary artery disease, have reported similar limitations of FRS and WHO charts, while broader algorithms such as QRISK-based models tend to identify a higher proportion of individuals as high risk [10]. The present study extends these observations by demonstrating comparable discrepancies across the glycemic spectrum, including individuals with prediabetes [3]. These findings highlight the importance of careful interpretation of cardiovascular risk estimates in Indian populations and support the need for population-specific validation of commonly used cardiovascular risk prediction models.

Strengths and Limitations

A key strength of this study is the inclusion of three distinct glycemic categories, allowing assessment of cardiovascular risk stratification across the continuum from normoglycemia to overt diabetes. The simultaneous application of four widely used cardiovascular risk prediction models within the same cohort minimizes inter-study variability and enables direct comparison [7,10]. Additionally, the Clinical Utility Index provided an additional framework for comparing classification behaviour across risk prediction models.

However, several limitations merit consideration. The single-center, cross-sectional design limits generalizability and precludes validation of predicted risk against observed cardiovascular outcomes. The absence of longitudinal cardiovascular outcomes limits assessment of predictive validity, and findings should be interpreted as comparative classification patterns. The modest sample size, particularly in the prediabetic group, may limit precision of subgroup analyses. These limitations underscore the exploratory nature of the findings.

The absence of a gold-standard cardiovascular outcome measure limited the ability to assess true predictive performance and restricted interpretation to comparative classification behaviour across models.

An additional methodological limitation is that diabetes mellitus was used as the surrogate reference classification state for comparative ROC analysis, while diabetes status also contributes directly to all evaluated cardiovascular risk prediction models. This may introduce partial circularity and overestimate relative classification concordance. Accordingly, ROC-derived measures should be interpreted only as comparative descriptors of classification behaviour rather than independent measures of diagnostic or prognostic performance.

Future Directions

The marked discordance observed among cardiovascular risk prediction models highlights the urgent need for population-specific validation and recalibration of existing tools or the development of indigenous risk prediction models tailored to Indian populations. Future studies should adopt prospective, longitudinal designs with hard cardiovascular endpoints to evaluate discrimination, calibration, and clinical impact. Future prospective investigations should additionally evaluate agreement among cardiovascular risk prediction models using pairwise cross-classification tables, Cohen's kappa statistics, weighted agreement coefficients, and reclassification analyses to provide a more comprehensive assessment of inter-model concordance. Incorporation of subclinical atherosclerosis markers and cost-effectiveness analyses may further inform optimal risk assessment strategies for resource-constrained settings.

Conclusion

This study demonstrates substantial variability in cardiovascular risk classification across commonly used prediction models when applied to Indian individuals across different glycemic states. These findings reflect differences in classification behaviour rather than predictive performance and highlight the need for population-specific validation using longitudinal cardiovascular outcome studies.

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

Author Contributions (CRediT):

LG.: Conceptualization, Investigation, Data curation, Formal analysis, Writing – original draft.
PR: Methodology, Data curation,
NT: Methodology, Data curation,
MKP: Methodology, Data curation,
JG: Formal analysis, Methodology, Data curation, Validation.
SNK: Conceptualization, Supervision, Writing – review & editing

Financial support and sponsorship

Funds received from JSSAHER to carry out the project.

Conflict of interest

The authors have no competing interests to declare that are relevant to the content of this article.

Data availability statement

The data underlying this article will be made available upon reasonable request to the corresponding author.

Ethical Clearance

Ethical Approval: The Institutional Ethics Committee clearance was obtained via letter no. Ref: JSS/MC/PG/16,17/2022-23.
Ethics Committee Name: JSS Medical College and Hospital Institutional Ethics Committee
Institution Name: JSS Medical College and Hospital,
Written informed consent was obtained from all the subjects for participation in the study and for the use of patient data for research and educational purposes. The study procedures

adhered to the guidelines of the Declaration of Helsinki 1964 and its subsequent revisions.

The manuscript has been read and approved by all the authors, and the requirements for authorship as stated earlier in this document have been met, and each author believes that the manuscript represents honest work.

References

1. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the Global Burden of Disease Study 2019. Journal of the American College of Cardiology. 2020;76(25):2982–3021. doi:10.1016/j.jacc.2020.11.010
2. Low Wang CC, Hess CN, Hiatt WR, Goldfine AB. Clinical update: cardiovascular disease in diabetes mellitus. Circulation. 2016;133(24):2459–2502.
3. Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ. 2016;355:i5953. doi:10.1136/bmj.i5953
4. Misra A, Khurana L. The metabolic syndrome in South Asians: epidemiology, determinants, and prevention. Metab Syndr Relat Disord. 2009;7(6):497–514.
5. Enas EA, Dhawan J, Petkar S. Coronary artery disease in Asian Indians: lessons learnt and the role of lipoprotein(a). Indian Heart J. 1997;49(1):25–34.
6. Kanjilal S, Rao V, Mukherjee M, Natesha B, Renuka K, Sibi K, et al. Application of cardiovascular disease risk prediction models and the relevance of novel biomarkers to risk stratification in Asian Indians. Vasc Health Risk Manag. 2008;4(1):199–211.
7. Lloyd-Jones DM, Braun LT, Ndumele CE, Smith SC Jr, Sperling LS, Virani SS, et al. Use of risk assessment tools to guide decision-making in the primary prevention of atherosclerotic cardiovascular disease. Circulation. 2019;139(25):e1162–e1177. doi:10.1161/CIR.0000000000000638.
8. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham heart study. Circulation. 2008;117(6):743–753.
9. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014;63(25):2935–2959.
10. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357:j2099. doi:10.1136/bmj.j2099
11. Bansal M, Kasliwal RR, Trehan N. Comparative accuracy of different risk scores in assessing cardiovascular risk in Indians: a study in patients with first myocardial infarction. Indian Heart J. 2014;66(6):580–586.
12. World Health Organization. Prevention of cardiovascular disease: guidelines for assessment and management of cardiovascular risk [Internet]. Geneva: World Health Organization; 2007. Available from: https://www.who.int/publications/i/item/9789241547178 (accessed: 31/12/2025)
13. Ridker PM, Cook NR. Statins: new American guidelines for prevention of cardiovascular disease. Lancet. 2013;382(9907):1762–1765.
14. Bansal M, Mehrotra R, Kasliwal RR. Cardiovascular risk stratification in Indians. J Clin Prev Cardiol. 2015;4(1):7–17 [Internet]. Available from: https://www.jcpcarchives.org/full/cardiovascular-risk-stratification-in-indians-153.php (accessed: 01/01/2026)
15. Doddamani P, Ramanathan N, Swetha NK, Suma MN. Comparative assessment of ADA, IDRS, and FINDRISC in predicting prediabetes and diabetes mellitus in South Indian population. J Lab Physicians. 2021;13(1):36–43.
16. Mitchell AJ. Sensitivity × PPV is a recognized test called the clinical utility index (CUI+). Eur J Epidemiol. 2011;26(3):251–252.
17. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29–36.
18. Qiu, P. (2005). The Statistical Evaluation of Medical Tests for Classification and Prediction. Journal of the American Statistical Association, 100(470), 705. https://doi.org/10.1198/jasa.2005.s19

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 application of indirect methods to establish and verify Argentinian paediatric thyroid function test reference intervals

Veronica E. Zaidman[1*], Melina B. Pelanda[1], Noelia Dujovne[2], Natalia Gazek[2], Claudia Insua[3], Rodolfo E. Recalde[1], Florencia Zabaleta[1], Eduardo A. Chaler[4], Juan M. Lazzati[1]

[1]*Laboratory of Endocrinology, J.P. Garrahan Pediatric Hospital, Ciudad Autónoma de Buenos Aires, Argentina*
[2]*Endocrinology department, J.P. Garrahan Pediatric Hospital, Ciudad Autónoma de Buenos Aires, Argentina*
[3]*Laboratory of Endocrinology, Pedro de Elizalde Hospital, Ciudad Autónoma de Buenos Aires, Argentina*
[4]*Laboratory consultancy, J.P. Garrahan Pediatric Hospital, Ciudad Autónoma de Buenos Aires, Argentina*

Article Info

[*]***Corresponding Author:***
Veronica Ethel Zaidman
Laboratory of Endocrinology, J.P. Garrahan Paediatric Hospital, Combate de los Pozos 1881, Ciudad Autónoma de Buenos Aires, Argentina. C.P.: 1245.
E-mail: vez2202@gmail.com

Keywords
thyroid function, reference ranges, paediatrics patients, indirect method, thyroid hormones

Abstract

The aim of this study was to establish and verify reference intervals for thyroid function tests using indirect methods in a paediatric hospital in Argentina. TSH and thyroid hormones were collected from a total of 19,822 patients. TSH, TT3 y FT4 were measured by Architect i4000 (Abbott) and TT4 by Immulite 2000 XPi (Siemens). Following the implementation of a series of filters to identify pathologies affecting the thyroid profile, a total of 5,871 cases remained in the database. Group A: 0-12 months (n=551). Group B: 13 months to 7 years old (n=2357). Group C: 8 to 18 years old (n=2963). Reference intervals were calculated by Hoffmann-Katayev and by CLSI28-A3. Verification was performed using the CLSI28-A3 protocol, as well as evaluated by being compared with the RCV. Reference intervals for TSH and thyroid hormones were established using Hoffmann–Katayev and CLSI28-A3 methods, showing overall agreement. CLSI28-A3 verification confirmed RIs for TSH and TT4 across all groups. For FT4 and TT3, only CLSI-derived RIs were consistently verified, while HK showed limited performance in younger groups. Comparison with published data using RCV demonstrated overall acceptable agreement, with some deviations in early ages. According to the verification procedure, suitable reference intervals calculation methods were Hoffmann–Katayev for TSH and TT4, and CLSI28-A3 for FT4 and TT3.Having reference intervals established for our patient population enhances the quality of results, improves their interpretation and safeguards patient safety. Indirect methods found to be good alternatives when a healthy population is not easily accessible.

Introduction

A test result in isolation, without accompanying information, is of limited use and fails to provide meaningful insights to the attending physician. Accurate interpretation requires the specification of units, methodology, brand of the assay and an appropriate reference interval (RI).

The concept of RI encompasses a substantial proportion of the reference population, the central 95% of the distribution, between $P_{2.5}$ and $P_{97.5}$. This leaves 5% of the reference population outside this range. There are four methods accepted by ISO 15189 norms for establishing RIs, in increasing order of strength: direct methods, indirect methods, bibliographic references, and the RI provided by the manufacturer [1,2]. All these RIs must be verified in the population to which they are to be applied.

Thyroid function is critical in several processes throughout life, especially in paediatrics. Depending on the severity and age of thyroid disease onset, early neurocognitive development and growth can be seriously affected [3]. Laboratory tests are a fundamental tool in diagnosing and monitoring thyroid disease. In addition to newborn screening, routine paediatric thyroid function screening is not clearly defined in current clinical care guidelines. The 2018 guide by the European Thyroid Association recommends screening for central hypothyroidism by testing for free thyroxine (FT4) and thyrotropin (TSH) in children with familial antecedents and/or failure to thrive, developmental delay, growth hormone (GH) deficiency, delayed or precocious puberty, or other hypothalamic–pituitary defects or lesions [3,4]. In previous research conducted by our group, the effectiveness of a full thyroid profile versus TSH plus FT4 in screening for thyroid disorders in paediatric patients was compared. The combination of TSH and FT4 has been demonstrated to be an effective method for the screening of thyroid disease in paediatric patients [5].

Therefore, TSH and thyroid hormones (THs) are the most frequently requested tests in the endocrinology laboratory of a high-complexity paediatric hospital for the evaluation and follow-up of thyroid disorders. They represent approximately 49% of all requests. TSH and free thyroxine (FT4) are particularly important, as they are essential for screening for thyroid disease in children and adolescents [5].

TSH and THs demonstrate significant variability, primarily due to age and, according to some authors, due to sex after puberty, as well as methodology and the specific brand of reagents used. Furthermore, there is a large amount of existing biological variation and inter-population differences [6]. Evident differences in methodological shows, even within the same methodology, among various reagents brands are highlighted in external quality assessment reports.

Specific populations exhibit distinct thyroid homeostasis. For instance, obese children have been found to have significantly higher TSH values than non-obese children [7-9]. One study demonstrated that substantial weight loss led to a significant decrease in TSH levels without the need for thyroid-specific treatment [8].

Consequently, it is essential for each laboratory to determine its own RIs, according to the population to be applied. Meanwhile, the establishment of an RI using direct methods is particularly problematic. The establishment of laboratory-specific direct RIs requires significant investment in reagents and specialized human resources, which is a limiting factor for several laboratories. Direct methods involve selecting a representative number of reference subjects, typically a minimum of 120 for each age and sex group, depending on the specific analyte requirement. Additionally, the availability of suitable reference individuals and the quantity of blood sample required are both significant considerations, particularly in a tertiary care paediatric hospital

In order to achieve the required number of individuals for each age/sex partition, a healthy volunteer program involving an interdisciplinary team is necessary. The establishment of RIs by direct methods for several analytes, including TSH and THs, was achieved through the CALIPER programme by Bailey et al [10]. They found a high variance in THs reference interval at birth, which significantly decreases by the first year of life and then remains constant. Absolute concentrations decline with age, except for total thyroxine (TT4), which decreases during childhood and surges between 14 and 18 years. Significant sex differences were also found in the adolescence group for TT4 and triiodothyronine (TT3). Regarding TSH, they demonstrate high variance at birth, this variance decreases gradually as the analyte levels increase with increasing postnatal age. These findings served as reference to constitute partitioning groups, in that study [10].

An alternative to direct methods is to use indirect ones, with a large data pool base of results. It has been proposed by some authors that the analysis of thousands of results, including those from pathological subjects, is statistically more robust than that of only 120 assuming healthiness subjects [11].

Indirect methods constitute a variant of data mining whereby a substantial database undergoes rigorous filtration, resulting in considerably lower costs when compared with direct methods. Hoffmann indirect method, which was developed in 1963, is the most widely applied [12,13]. Hoffmann assumed that the data would follow a Gaussian or near-Gaussian distribution, but this is a significant limitation in the field of endocrinology. The majority of hormones do not have a normal distribution.

In 2010, Katayev put forward a computerized modification to the Hoffmann method [11]. The present modification is an attempt to reduce the subjectivity associated with RI calculations and enhance reproducibility by establishing a logical progression. After a thorough and meticulous database filtration procedure, outliers are eliminated. Following this, the data is analysed using a cumulative frequency chart.

The methodology entails the iterative removal of the distal portions of the graph until a satisfactory linear approximation is attained. The RI is determined through the extrapolation of the 2.5th and 97.5th percentiles of the cumulative frequency plot on

the concentration axis.
The calculated RI must undergo verification in order to confirm its validity and reliability. Katayev proposes the implementation of reference change values (RCV) as a statistical significance indicator, in comparison with RI as previously documented in the bibliography [11].
The Clinical and Laboratory Standards Institute (CLSI) has published a recommendation for a RI verification algorithm in the CLSI28-A3 guide [14]. This algorithm involves the recruitment of 20 subjects selected from the target population, who are required to be free of any underlying conditions that could affect the analyte. In instances where no more than two patients fall outside the RI, the RI is considered confirmed and suitable for clinical application. In the case of three or four results falling outside a second verification step, identical to the first, this step is repeated. When the number of individuals exceeding the RI limits is five or more, the RI is not appropriate to the target population [15].

Objective

The aim of this study was to establish and verify reference intervals for thyroid function tests using indirect methods in a paediatric tertiary care hospital in Argentina.

Material and methods

Assay Methods

TSH, TT3 and FT4 were performed using chemiluminescent microparticle immunoassay (CMIA) by Architect i4000 (Abbott Ireland Diagnostics Division, Longford, Ireland), while TT4 was tested using chemiluminescent immunoassay by IMMULITE© 2000XPi (Siemens Healthineers, Llanberis, UK). The serum samples collected in gel tubes were processed on the same day as their collection.

Database

The data for TSH and THs were collected from 19,822 patients of J.P. Garrahan Paediatric Hospital between January 2020 and December 2021. In cases where multiple values were available for the same analyte in the same patient, only the oldest result was selected for inclusion in the analysis.
In order to avoid the inclusion of patients affected by non-thyroidal illness syndrome, all inpatients were excluded from the database. Exclusion criteria included prior diagnosis of thyroid disease (thyroiditis, thyroid cancer, hypothyroidism, or hyperthyroidism) and previous thyroidectomy. The general laboratory data available was subjected to evaluation in order to identify patients with possible abnormal thyroid profiles. This was achieved by filtering out patients with abnormal liver function tests and/or a pathological hematology panel, as well as those with positive anti-thyroid antibodies results, using as positive the manufacturer-reported cutoff values of anti-thyroid peroxidase antibodies >35 IU/ml and/or anti-thyroglobulin antibodies >40 IU/ml [16, 17]. The reference intervals applied for filtering in this study, for both the haematology panel (haematocrit and haemoglobin) [18, 19] and the liver function panel [20] were those previously established and verified for routine clinical use at our institution. Following the curation of the initial 19,822 patients' data, a total of 5,871 remained in the database for the purposes of further analysis.

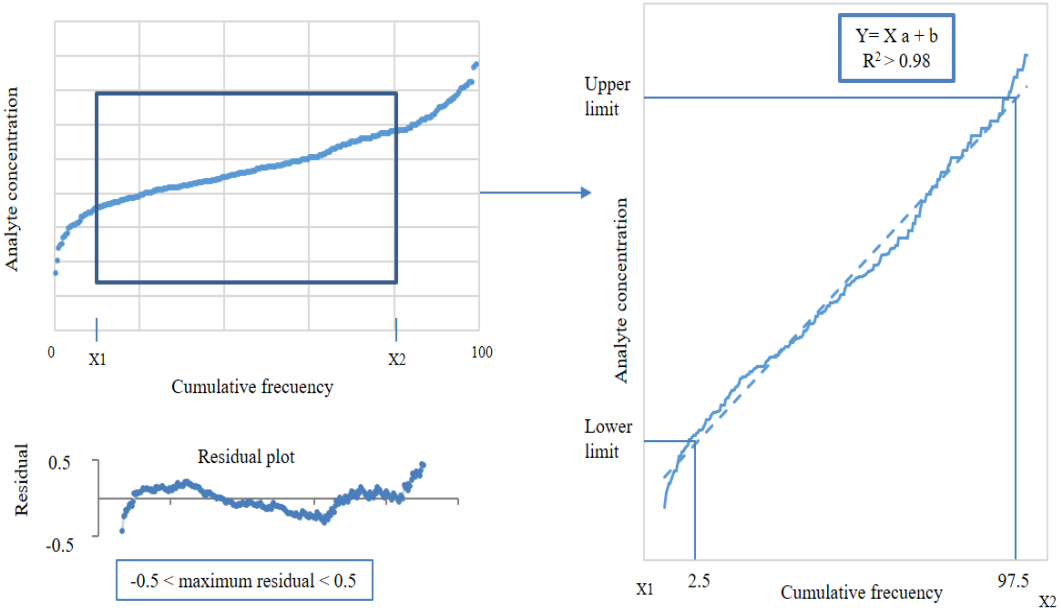
Partitions group

A total of 5871 patients were divided into three groups for the purposes of the study. Group A comprised infants aged 12 months and under (n = 551, female: 43.4%). Group B included children aged 13 months to 7 years old, inclusive (n =2357, female: 49.9%). Group C contained children from 8 to18 years old (n = 2963, female: 56.9%).

Indirect methods analysis

RIs were calculated in each group applying two methods. Firstly, the Hoffmann-Katayev adapted method (HK) was employed [11]. This method involved graphing cumulative frequency versus test value without transformation. Outliers were identified using Tukey's method and were excluded [21]. Then, the extremes were trimmed to improve the linear regression within an acceptable margin of error. The following rules were applied: A $r^2 \geq 0.98$ was considered adequate and the maximum absolute residual do not exceed 0.5. Upon fulfillment of both criteria, the trimming process was halted (Figure 1). After that, the RIs were obtained by extrapolating the values from the linear regression of the 2.5th and 97.5th percentiles.
In the second method, RIs were calculated using percentiles according to the procedure recommended for non-parametric analytes by the Clinical and Laboratory Standards Institute (CLSI) (CLSI28-A3) [14].

Figure 1: Calculation scheme for RIs using the Hoffmann-Katayev method.



RI verification

A total of 20 independent serum samples from children corresponding to each age group with presumably normal thyroid function, who had undergone pre-surgical laboratory testing, were used to verify the calculated RI. A specialist was assigned to review the subjects' clinical records with a view to excluding any pathologies or medications that might have an impact on thyroid function. The presence of outliers was detected and subsequently excluded by Tukey method [21]. The processing of the data was conducted in accordance with the guidelines stipulated in CLSI EP28-A3c [14].

Assessment of clinical relevance (comparison test with RCV)

In addition, as a comparative exploration, the calculated RIs were compared to those previously published by Bailey et al. as part of the CALIPER project. (CALG) [10] and by Bokulić et al. (CROG) [22] both were obtained by direct method. The percentage of difference was calculated for each comparison and then compared to the RCV defined as: RCV = $\sqrt{2}$ x Z x $\sqrt{(CVa^2 + CVi^2)}$. Where Z is the probability selected by significance is 1.96 for p=0.05. CVa is the inter-assay coefficient of variation and CVi within–subject biological variation from the biological variation table. For this purpose, the RCV was used as the maximum accepted variation when comparing the lower and upper limits of the RI of each group and calculation methodology.

In instances where the comparison age groups do not correspond to those that have been established in this study, a comparative analysis was conducted with all RIs that include age groups G1, G2, and G3.

Results

Obtained TSH and THs RIs are shown in Table 1. The percentage of values trimmed to obtain the RIs using the HK method was as follows: TSH: 14.6%, 15.2%, and 10.3%; FT4: 8.4%, 10.0%, and 7.7%; TT3: 5.7%, 12.5%, and 8.9%; TT4: 18.0%, 14.3%, and 9.4%, for Groups A, B, and C, respectively. Both methods yielded similar RIs, lower and upper limits showed a difference of less than 20%, except for TSH at upper limit in group A and B, where significant differences were observed, although the RCV was exceeded only in group A (Table 1). Calculated upper limits by HK were lower than those obtained by CLSI28-A3, the opposite occurred with lower limits (Table 1).

Table 1: Reference interval calculated by indirect methods.

	group	Hoffmann-Katayev		CLSI28A3 (90% confidence interval)		Relative difference % Hoffmann-Katayev vs CLSI28A3	
		Lower limit	Upper limit	Lower limit	Upper limit	Lower limit	Upper limit
TSH (µUI/ml)							
0 - 12 m	A	0.68	4.48	0.56 (0.51-0.70)	8.26 (7.50-8.65)	19.4	-59.3
13 m - 7 y	B	0.85	4.84	0.80 (0.75-0.87)	6.15 (5.93-6.34)	6.1	-23.8
8 - 18 y	C	0.76	4.27	0.72 (0.69-0.79)	5.17 (5.09-5.26)	5.4	-19.1
FT4 (pmol/l)							
0 - 12 m	A	10.81	15.83	10.42 (10.04-10.68)	17.25 (16.60-17.50)	3.7	-8.6
13 m - 7 y	B	11.33	15.06	10.04 (9.78-10.30)	17.25 (16.99-17.37)	12.0	-13.5
8 - 18 y	C	10.55	14.29	9.91 (9.90-10.04)	15.96 (15.83-16.09)	6.3	-11.1
TT3 (nmol/l)							
0 - 12 m	A	1.84	3.56	1.70 (1.60-1.75)	4.07 (3.92-4.16)	7.8	-13.3
13 m - 7 y	B	1.64	3.12	1.52 (1.49-1.55)	3.59 (3.52-3.61)	7.8	-14.2
8 - 18 y	C	1.41	2.73	1.17 (1.14-1.27)	3.04 (2.98-3.07)	19.0	-10.6
TT4 (nmol/l)							
0 - 12 m	A	84.94	162.16	78.51 (74.65-82.37)	191.76 (184.04-196.91)	7.9	-16.7
13 m - 7 y	B	83.66	146.72	72.07 (70.40-73.10)	158.30 (154.44-160.88)	14.9	-7.6
8 - 18 y	C	77.22	132.56	65.64 (64.86-66.92)	142.86 (140.28-144.14)	16.2	-7.5

Verification using the CLSI protocol showed that TSH and TT4 were approved for all groups, calculated HK, CLSI28-A3 RIs and previously published RIs. In the case of FT4, the RI calculated by CLSI28-A3 and those previously published by CALG [10] and CROG [22] were verified for all groups. However, the RI calculated by HK was not verified for any group. For TT3, the RIs calculated by CLSI28-A3 were verified for all groups. However, this was not the case for RIs calculated by HK in G1 and G2. For the G3 group, all RIs were verified. When examining the TSH values calculated by HK RIs using RCV, an acceptable percentage deviation was found compared to the published values by CALG and CROG. In the case of FT4, the RIs calculated using both methods were comparable with CROG, with the exception of G1, where the percentage difference compared to the group was greater, in the following age group 2 to 15 days.

In the instance of TT3, the RIs calculated by HK were acceptable in G2 and G3. However, when we compared the RIs obtained by CLSI28-A3 with the published values, some differences exceeded the RCV%, as is the case with FT4 in G1, where the difference was greater compared to the group from 2 to 15 days of the CROG and were unsuccessful when comparing with CALG, TSH in G1 and TT3 in G2, where there is a disagreement with the group from 6-14 of the CROG that overlaps only one year with G3 group (Table 2). All of the RIs comparisons were successful for TT4 when compared with the manufacturer's published values.

Table 2: Comparison test with RCV: differences compared with the high and low levels reported in the literature.

Group	Analite	CROG					CALG				
		ARG	Lower limit		Upper limit		ARG	Lower limit		Upper limit	
			% HK	% CLSI 28-A3	% HK	% CLSI 28-A3		% HK	% CLSI 28-A3	% HK	% CLSI 28-A3
G1 0-12m	TSH	2d-<2m	30.5	11.3	-37.4	23.2	4d-<6m	-7.1	-26.4	-6.3	53.6
		2m-<14y	-16.2	-35.3	-2.9	56.7*	6m-<14y	-2.9	-22.2	7.2	65.8*
	TT3	2d-<1y	59.5*	52.3*	-0.9	12.4	4d-<1y	35.3*	27.7*	-0.9	12.4
	FT4	2d-<15d	-9.1	-12.7	-55.3*	-47.3*	5d-<15d	-22.2	-25.8*	-89.2*	-82.2*
		15d-<1y	-4.7	-8.3	-17.8	-9.2	15d-<30d	21.1	17.4	-69.1*	-61.5*
							30d-<1y	-5.8	-9.4	-32.1*	-23.7*
G2 13m-7y	TSH	2m-<14y	6.1	0.0	4.9	28.6	6m-<14y	19.4	13.3	14.9	38.4
	TT3	1y-<6y	-2.8	-10.5	7.7	21.8	1y-<12y	-5.5	-13.2	7.1	21.3
		6y-<14y	16.2	8.4	10.9	25.0*					
	FT4	1y-<19y	8.3	-3.8	-5.0	8.6	1y-<14y	-1.1	-13.2	-15.7	-2.2
G3 8-18y	TSH	2m-<14y	8.2	2.8	-7.7	11.5	6m-<14y	8.2	2.8	2.4	21.4
		14y-<19y	47.2	42.0	17.0	35.8	14-<19y	47.2	42.0	22.4	41.0
	TT3	6-<14y	1.1	-18.0	-2.2	8.4	1-<12y	-20.5	-39.2*	-6.0	4.7
		14-<19y	7.9	-11.2	17.1	27.6*	12-<15y	-6.3	-25.3*	1.1	11.8
	FT4	1y-<19y	1.2	-5.1	-10.3	0.8	1-<19y	-8.2	-14.5	-21.0	-10.0

*RCV: TSH:55.5%, FT4:21.4% and TT3:23.5%. ARG: Age range of the comparison group.HK: Hoffmann Katayev method. *The difference exceeds the RCV.*
CALG: Bailey et al. [10] and CROG: Bokulić et al. [22].

In accordance with the findings derived from the verification process, Table 3 presents the most suitable TSH and THs RIs for the paediatric population under study.

Table 3: Reference intervals for TSH and thyroid hormones selected on the basis of the verification results to be applied.

	Lower limit	Upper limit	Source
TSH (µUI/ml)			
0 - 12 m	0.68	4.48	Hoffmann-Katayev
13 m - 7 y	0.85	4.84	
8 - 18 y	0.76	4.27	
FT4 (pmol/l)			
0 - 12 m	10.42	17.25	CLSI28-A3
13 m - 7 y	10.04	17.25	
8 - 18 y	9.91	15.96	
TT3 (nmol/l)			
0 - 12 m	1.70	4.07	CLSI28-A3
13 m - 7 y	1.52	3.59	
8 - 18 y	1.17	3.04	
TT4 (nmol/l)			
0 - 12 m	84.94	162.16	Hoffmann-Katayev
13 m - 7 y	83.66	146.72	
8 - 18 y	77.22	132.56	

Discussion

Obtaining specific RIs for the target population with the available equipment represents a constant challenge for paediatric endocrinology laboratories. In this study, RIs were obtained using the indirect method with two alternatives, Hoffmann and CLSI. An accessible alternative was found, as it does not require a large investment in human or economic resources.

A large database comprising 19,822 patients was analysed, and following a rigorous filtration process, this was reduced to 5,871. The large number of patients, presumed to be free of thyroid disease, lends strength to the RIs obtained. In addition, the processing of data over a two-year period mitigates seasonal variations, changes in reagent batches and sampling materials, consequently reducing pre-analytical and analytical variations.

The intervals obtained were verified to determine whether they could be used for our population. To this end, the CLSI-recommended method was used, which yielded favourable results for all RIs obtained for TSH and TT4. For TT3 and FT4, the RIs calculated by CLSI28-A3 were verified, but not those calculated by HK, as they were found to be very narrow. This outcome can be attributed to the calculation procedure, which was developed with the objective of minimizing error.

Furthermore, in order to evaluate the RIs obtained, the RCV was applied as a parameter to test the clinical relevance of the RIs obtained with respect to those previously published in the literature, calculated by the direct method, determining the percentage difference. With regard to TSH, the results obtained were favourable, with the exception of the upper limit in G1. In this case, the publications utilized for comparative analysis implemented a distinct categorization of age ranges, including patients up to 14 years of age within the same category as those below 1 year of age, and their upper limits were considerably lower than our RI calculated by CLSI28-A3; the one calculated by HK was verified. In addition, TSH exhibits greater variability, with a RCV of 55.5%, which is reflected in the presence of significant differences at the upper limit of the reference range. Numerous publications have documented the phenomenon of TSH variability, even when the same measuring method or analytical platform is analysed.

Analysing the FT4 data, it was found that the comparison test with RCV was successful in comparison to CROG, with the exception of infants younger than 15 days, as this group was not separated during the process of this study. In TT3, the comparison test with RCV demonstrated that the RIs calculated by HK exhibited less discrepancy with those previously published than those calculated by CLSI28-A3.

The principal aim of this study is summarized in Table 3, which details the RIs that were found to be appropriate for the study population. These RIs were based on the results obtained using the CLSI recommended verification method and the comparison test with RCV.

Limitations

Obtaining a processable database from a tertiary care institution's system proved to be challenging. The filtering of a large database, although restrictive, may still yield a set of abnormal values; however, this is something that the indirect method already takes into account in its calculation. However, certain parameters must be established to standardize the calculation, such as the method for excluding outliers or the acceptable error in the linearity of the HK curve, which determines how far to trim the tails. It is important to establish a method that ensures the RI calculation is not dependent on the operator's experience.

In order to verify the RIs using CLSI protocol, we encountered significant challenges in obtaining samples from children without any medical conditions or treatments that might affect their thyroid profile. This was due to the fact that the hospital is a tertiary care facility and the volume of paediatric samples is limited.

The sample size within the under-1-year-old group was insufficient to permit further subgroup stratification; this resulted in significant differences, particularly in FT4, when compared with previously published data from CALG and CROG, which provide reference intervals for infants under 15 days of age. Certain precautions should be taken when using the reference intervals published in this study for infants younger than 15 days. Future studies will focus on this age group in order to reach the number of subjects required to allow further subdivision, especially within the 0–12-month range, where neonatal and early infancy thyroid physiology changes rapidly.

The recruitment of 20 healthy subjects for Group A to the verification process proved to be a significant challenge. Apparently, endocrine-healthy children under 13 months of age do not require laboratory analysis. In order to achieve the necessary sample size, assistance was requested from a nearby paediatric hospital, Elizalde Hospital. The samples were also processed on the same equipment used to obtain the results for calculating the reference intervals (RIs). Despite the similarity in terms of demographics between the two patient populations, the conditions treated at Elizalde Hospital are less complex. This constitutes a limitation of the study, as identical patient populations were not used to verify the RIs for Group A, since it was not possible to obtain samples from infants under one year of age with no apparent thyroid condition.

Significant differences in THs in adolescents by biological sex. have been shown in previous studies, such as those published by Bailey et al. [10] and by Campbell et al [23]. To investigate further, a Mann-Whitney test was conducted, which revealed significant differences between males and females in the 12 to 18 age group for both TT3 and TT4. Due to the design of the database, distinguishing between the two groups to obtain the RIs is beyond the scope of this study.

Conclusions

Having reference intervals established for our patient population, under the same pre-analytical and analytical conditions as those used in our daily work, enhances the quality of results, improves their interpretation and safeguards patient safety.

Thus, indirect methods have been found to be good alternatives when a healthy population is not easily accessible. However, since these methods involve some subjective decision points, it is important to carry out a thorough verification before they are applied.

Author Contributions

Veronica E. Zaidman: Conceptualization, methodology, validation, formal analysis, investigation, resources, data curation, writing - original draft, writing - review & editing, visualization and project administration.

Melina B. Pelanda: Validation, investigation, resources and writing - review & editing.

Noelia Dujovne: Investigation, resources and writing - review & editing.

Natalia Gazek: Investigation, resources and writing - review & editing.

Claudia Insua: Investigation, resources and writing - review & editing.

Rodolfo E. Recalde: Validation, investigation and writing - review & editing.

Florencia Zabaleta: Validation, investigation and writing - review & editing.

Eduardo A. Chaler: Conceptualization, methodology, formal analysis, investigation, resources, data curation, writing - review &editing, supervision and project administration.

Juan M. Lazzati: Conceptualization, methodology, investigation, resources, data curation, writing - review &editing, supervision and project administration.

Author's Disclosures

The authors declare no conflicts of interest for the conduction of the study. The corresponding author is responsible for ensuring that the descriptions are accurate and agreed upon by all authors.

Ethics Approval

This study utilised an anonymised patient database. Consequently, it did not require approval from the ethics committee. The research was conducted in accordance with the Declaration of Helsinki and national and institutional standards.

Data Availability Statement

The databases analysed during the current study are not publicly available but are available from the corresponding author on reasonable request.

IA Statement

During the preparation of this work the authors used DeepL IA in order to improve grammar and clarify language. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the published article.

Funding Statement

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

References

1. Chaler EA. Intervalos de referencia en pediatría. Una guía de sus formas de cálculo ordenadas según su fortaleza metodológica. Acta BioquímClínLatinoam. 2022;2957(1):75–80. ISSN 1851-6114.
2. Dirks NF, Den Elzen WPJ, Hillebrand JJ, Jansen HI, Boekel E Ten, Brinkman J, et al. Should we depend on reference intervals from manufacturer package inserts? Comparing TSH and FT4 reference intervals from four manufacturers with results from modern indirect methods and the direct method. ClinChem Lab Med [Internet]. 2024;1–10. https://doi.org/10.1515/cclm-2023-1237.
3. Hanley P, Lord K, Bauer AJ. Thyroid disorders in children and adolescents: A review. JAMA Pediatr. 2016;170(10):1008–1019. DOI: 10.1001/jamapediatrics.2016.0486.
4. Persani L, Brabant G, Dattani M, Bonomi M, Feldt-Rasmussen U, Fliers E, et al. 2018 European Thyroid Association (ETA) Guidelines on the Diagnosis and Management of Central Hypothyroidism. Eur Thyroid J. 2018;7(5):225–237. DOI: 10.1159/000491388.
5. Zaidman V, Lazzati JM, Maceiras M, Herzovich V, Pelanda M, Ribas A, Belgorosky A. and Chaler E. Rational approach to the primary evaluation of thyroid disease in paediatrics. Full thyroid profile vs. thyroid-stimulating hormone and free thyroxine only. ClinChem Lab Med [Internet].2018; 5–7. https://doi.org/10.1515/cclm-2017-0962.
6. Ankrah-Tetteh T, Wijeratne S, Swaminathan R. Intraindividual variation in serum thyroid hormones, parathyroid hormone and insulin-like growth factor-1. Ann ClinBiochem 2008; 45:167–169. DOI: 10.1258/acb.2007.007103.
7. Ghergherehchi R, Hazhir N. Thyroid hormonal status among children with obesity. TherAdvEndocrinolMetab. 2015;6(2):51–55. DOI: 10.1177/2042018815571892.
8. Lass N, Barth A, Reinehr T. Thyroid Volume and Thyroid Function Parameters Are Independently Associated with Weight Status in Overweight Children. Horm Res Paediatr. 2020;93(5):279–286. DOI: 10.1159/000509786.
9. Walczak K, Sieminska L. Obesity and thyroid axis. Int J Environ Res Public Health. 2021;18(18). https://doi.org/10.3390/ijerph18189434.

10. Bailey D, Colantonio D, Kyriakopoulou L, Cohen AH, Chan MK, Armbruster D, et al. Marked biological variance in endocrine and biochemical markers in childhood: Establishment of pediatric reference intervals using healthy community children from the CALIPER cohort. Clin Chem. 2013;59(9):1393–1403. DOI:10.1373/clinchem.2013.204222.
11. Katayev A, Balciza C, Seccombe DW. Establishing Reference Intervals for Clinical Laboratory Test Results Is There a Better Way ? 2010;180–6. DOI: 10.1309/AJCPN5BMTSF1CDYP.
12. Hoffmann RG. Statistics in the Practice of Medicine. JAMA J Am Med Assoc. 1963;185(11):864–873. DOI: 10.1001/jama.1963.03060110068020.
13. Jones GRD, Haeckel R, Loh TP, Sikaris K, Streichert T, Katayev A, et al. Indirect methods for reference interval determination - Review and recommendations. ClinChem Lab Med. 2019;57(1):20–29. https://doi.org/10.1515/cclm-2018-0073.
14. Horowitz G, Altaie S, Boyd J, Ceriotti F, Garg U, Horn P, et al. Defining, establishing, and verifying reference intervals in the clinical laboratory. Clinical and Laboratory Standards institute. 2010. p. 1–37. ISBN 1-56238-682-4.
15. Ozarda Y, Higgins V, Adeli K. Verification of reference intervals in routine clinical laboratories: Practical challenges and recommendations. ClinChem Lab Med. 2019;57(1):30–37. https://doi.org/10.1515/cclm-2018-0059
16. Siemens Healthcare Diagnostics. IMMULITE 2000 Anti-TG Ab (PIL2KTG-22, 2015-08-04). Llanberis, UK: Siemens Healthcare Diagnostics Products Ltd.; 2015 Available from: https://www.siemens-healthineers.com
17. Siemens Healthcare Diagnostics. IMMULITE 2000 Anti-TPO Ab (PIL2KTO-24, 2018-03-15) . Llanberis, UK: Siemens Healthcare Diagnostics Products Ltd.; 2018. Available from: https://www.siemens-healthineers.com
18. Dallman PR. In: Rudolph A, editor. Pediatrics. 16th ed. New York: Appleton-Century-Crofts; 1977. p. 1111. Nathan and Oski´s, Hematology of Infancy and Childhood, 5th edition, 1998.
19. Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). National Health and Nutrition Examination Survey Data, 1999-2000. Hyattsville, MD: U.S. Department of Health and Human Services, CDC; 2001. Available from: https://www.cdc.gov/nchs/nhanes/nhanes1999-2000/nhanes99_00.htm
20. Heil W, Ehrhardt V. Reference Ranges for Adults and Children: Pre-Analytical Considerations. 9th ed. Mannheim: Roche Diagnostics GmbH; 2008. p. 14-16, 22-24, 46.
21. Beyer, H. Tukey, John W.: Exploratory Data Analysis. Addison-Wesley Publishing Company Reading, Mass.S.. Biom. J., 23: 413-414. https://doi.org/10.1002/bimj.4710230408.
22. Bokulić A, Zec I, Marijančević D, Goreta S, Šepec MP, Sabolić LLG. Establishing paediatric reference intervals for thyroid function tests in Croatian population on the Abbott Architect i2000. Biochem Med. 2021;31(3):030702. https://doi.org/10.11613/BM.2021.030702 .
23. Campbell PJ, Brown SJ, Kendrew P, Lewer M, Lim EM, Joseph J, Cross SM, Wright MJ, Martin NG, Wilson SG, Walsh JP. Changes in Thyroid Function Across Adolescence: A Longitudinal Study. J ClinEndocrinolMetab. 2020;105(4): dgz331. doi: 10.1210/clinem/dgz331. PMID: 31922575.

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

Multicomponent CAPA Programme Incorporating Cloud-Based Laboratory Information System Migration for Post-Analytical Error Reduction in an Accredited Stand-Alone Diagnostic Laboratory: A 13-Month Before-and-After Study

Jagdish Chandarana[1*], Kinjal Butani[1]

[1]*Vedant Diagnostic Center, Rajkot, Gujarat, India*

Article Info

[*]***Corresponding Author:***
Jagdish Chandarana
Vedant Diagnostic Center, Rajkot, Gujarat 360002, India
E-mail: jagdish@vedantdiagnostic.com
Phone: +91 98252 19517
ORCID: 0009-0002-8796-6432

Keywords

Post-analytical errors, Laboratory information system, Cloud computing, Quality improvement, Six Sigma, Medical informatics

Abstract

Purpose: Post-analytical errors are under-investigated relative to pre-analytical errors, particularly in stand-alone diagnostic laboratories in low- and middle-income countries (LMICs). We evaluated a multicomponent Corrective and Preventive Action (CAPA) programme — combining process-level controls with two infrastructure changes, migration from a local on-premise laboratory information system (LIS) to a centralised cloud-based LIS and installation of dual-broadband connectivity — for reducing post-analytical errors at an NABL-accredited diagnostic laboratory.

Methods: Single-centre before-and-after observational quality-improvement study reported per SQUIRE 2.0. Failure Mode and Effects Analysis (FMEA) identified six post-analytical failure modes. A six-month pre-CAPA baseline (April–September 2025; 19,898 reports) was compared with a six-month post-CAPA period (November 2025–April 2026; 20,794 reports). October 2025 was designated a priori as an implementation window and excluded from the primary comparison. Outcomes were error count, errors per million reports released (EPMR), and Sigma level. Exact Poisson rate ratios, Fisher's exact test and a bounding analysis for the excluded transition month were performed.

Results: Post-analytical errors fell from 24 (1,206 EPMR; ~4.53σ) to 4 (192 EPMR; ~5.05σ), an 84.1% reduction (rate ratio 0.159, 95% CI 0.040–0.464; Fisher's exact p = 0.00008). The 10 pre-CAPA LIS/software errors arose from two discrete infrastructure incidents, each affecting five reports; no such incident was recorded post-CAPA. Every error category fell. The LIS failure mode was correctly scored low risk on clinical severity and event frequency, yet generated the largest share of observed errors. Bounding analysis indicated the reduction would remain significant unless the excluded transition month contained ≥11 errors, exceeding the highest observed baseline month (7).

Conclusions: A multicomponent CAPA package produced a substantial post-analytical error reduction and a ~0.5 Sigma improvement. The reduction cannot be attributed to any single component. The finding that prospective FMEA classified the dominant failure mode as low risk supports complementing prospective risk assessment with retrospective error logging.

Introduction

Errors in clinical laboratory testing remain a persistent threat to patient safety, with consequences extending from specimen recollection and diagnostic delay to misdiagnosis, inappropriate therapy, and measurable patient harm [1]. The total testing process is conventionally divided into pre-analytical, analytical, and post-analytical phases, and the relative contribution of each to total error burden has been studied for more than two decades [2, 3]. While the pre-analytical phase has historically attracted the greatest attention, accounting for an estimated 60–70% of total errors in most large series [4], the post-analytical phase — result validation, interpretive review, critical value notification, turnaround time management, result transmission, and post-test sample handling — has received comparatively less systematic study [5].

Post-analytical errors are consequential because they occur at the laboratory–clinician interface where corrective opportunities are limited and clinical action is imminent [6, 7]. The IFCC Working Group on Laboratory Errors and Patient Safety has extended its harmonised quality-indicator framework to post-analytical indicators [8, 9], and the 2022 revision of ISO 15189 strengthened pre- and post-examination requirements [10]. Despite this, the evidence base is concentrated in hospital laboratories in high-income countries, with stand-alone diagnostic laboratories in LMIC settings under-represented [11].

Stand-alone diagnostic laboratories in LMIC contexts face distinct operational realities. Specimen collection is often distributed across collection centres, and infrastructure may rely on locally maintained servers rather than centrally managed enterprise information systems [22, 23]. Under such arrangements, server crashes, software hangs, and connectivity disruptions can be recurrent operational events affecting report release [11, 12, 33]. Laboratory information systems in resource-constrained settings frequently lack the redundancy, monitoring and vendor support taken for granted in enterprise environments, and implementation experience from such settings remains sparsely documented [33]. Structured CAPA frameworks anchored in NABL and ISO 15189 offer an institutional infrastructure for systematic improvement, but their application to the post-analytical phase in stand-alone LMIC laboratories is poorly documented [13].

Migration of clinical information systems to cloud-hosted architecture has been advanced as a route to improved availability and reduced local maintenance burden [24, 25], but it redistributes rather than eliminates risk: dependency shifts from local hardware to network availability, vendor service levels, and the security posture of the hosting provider [27, 34]. Health informatics scholarship further emphasises that clinical information systems are socio-technical rather than purely technical entities, and that their safety properties emerge from the interaction of software, workflow and the people who use them [26, 28]. Cybersecurity exposure in healthcare information systems has grown substantially [27], and health information technology implementations frequently fail for organisational rather than technical reasons, including user resistance and inadequate change management [29, 30, 31]. Evidence that health IT interventions improve quality is real but heterogeneous, and effects are strongly context-dependent [32]. These considerations are directly relevant when infrastructure change is proposed as a quality intervention.

Six Sigma methodology, expressed through defect rates and Sigma levels, has been widely adopted for analytical quality monitoring [14] and extended to the pre- and post-analytical phases [15, 16]. Recent work shows that resource-appropriate interventions in stand-alone NABL-accredited laboratories can achieve internationally benchmarkable pre-analytical Sigma performance [17], but corresponding post-analytical evidence is sparse. The present study reports a 13-month before-and-after evaluation of a multicomponent post-analytical CAPA programme combining process-level controls with two infrastructure changes. The aims were to (i) quantify the change in post-analytical error frequency and Sigma performance, (ii) describe the contribution of individual components without attributing the aggregate effect to any one of them, and (iii) compare prospective FMEA risk scoring with observed error frequencies.

Materials and Methods

This quality-improvement study is reported in accordance with the SQUIRE 2.0 guideline [21] (completed checklist provided as Supplementary Information).

Study setting

The study was conducted at a stand-alone clinical diagnostic laboratory in western India, accredited by the National Accreditation Board for Testing and Calibration Laboratories (NABL) under ISO 15189 and operating within a partner-laboratory network arrangement. Across the 13-month study period the laboratory released a mean of 3,393 patient reports per month (range 2,989–3,910) across clinical chemistry, haematology, immunology, serology and microbiology, serving an outpatient catchment through a network of collection centres. Monthly report volumes were extracted retrospectively from the laboratory information system and are reported in full in Table 1; they replace the approximate monthly figure used in earlier internal reporting.

Study design and definition of study periods

This was a single-centre observational before-and-after quality-improvement study spanning a 13-month calendar period (1

April 2025 to 30 April 2026), of which 12 months contributed to the primary analysis: a pre-CAPA baseline (1 April 2025 to 30 September 2025; six months; 19,898 reports) and a post-CAPA observation period (1 November 2025 to 30 April 2026; six months; 20,794 reports). A total of 40,692 patient reports were released across the two analysed periods.

October 2025 (3,419 reports) was designated a priori as an implementation window rather than an observation period, and was excluded from the primary comparison. The rationale was that the LIS migration was executed progressively across the whole month rather than as a single-day cutover; throughout October the information system was in neither its pre-intervention nor its post-intervention configuration, and errors arising during this state are attributable to neither. Consistent with this designation, the laboratory's error log was operated as a phase-locked study instrument and post-analytical errors were not counted during the transition month. Because this exclusion could in principle favour the intervention, a bounding analysis quantifying its maximum possible influence is reported in the Results and the absence of transition-month error counts is acknowledged as a limitation.

Failure Mode and Effects Analysis

Prior to study commencement the quality-improvement team conducted a structured FMEA of the post-analytical phase, adapted from established laboratory FMEA frameworks [18]. The assessment was performed in March 2025 by a four-member team comprising the managing director (quality management), the laboratory head/pathologist, the senior scientific officer, and the quality-improvement coordinator. Six post-analytical failure modes were identified from internal incident records, NABL non-conformity logs, and team discussion.

Each mode was scored on Severity (S) and Probability (P) using five-point scales. Severity anchors were: 1 = negligible, no clinical consequence; 2 = minor, administrative correction only; 3 = moderate, potential to affect a clinical decision; 4 = major, incorrect or delayed clinical action likely; 5 = critical, potential for patient harm. Probability anchors were expressed as expected event frequency: 1 = rare (<1 event/year); 2 = unlikely (1–4 events/year); 3 = possible (5–12 events/year); 4 = likely (13–24 events/year); 5 = frequent (>24 events/year). Risk Priority Number (RPN) was calculated as S × P, with categories Low (<5), Medium (5–8), High (9–14) and Extreme (≥15). Scoring was assigned by group consensus in a single facilitated session rather than by independent rating followed by reconciliation; disagreements were resolved by discussion until consensus was reached. Detectability was not incorporated, since several failure modes were by definition identified only after report release, making a detectability rating uninformative. No formal inter-rater reliability statistic was calculated. These are acknowledged as methodological limitations. The full scoring matrix and anchor definitions are provided as Supplementary Information. A formal post-intervention residual-risk reassessment was not undertaken; this is identified as a limitation.

We note explicitly that in the manuscript as originally drafted, two failure modes scoring RPN 4 (TAT Breach and LIS/Software Failure) were reported as Medium risk. Applied consistently, the categories defined above classify RPN 4 as Low risk. This has been corrected throughout; the correction strengthens rather than weakens the principal FMEA observation reported below.

Operational definitions and error adjudication

TAT Breach (Routine/STAT): any report not released within the laboratory's locally documented turnaround-time targets, attributable to post-analytical workflow factors. The laboratory operates under a franchise arrangement in which locally documented TAT targets, rather than the partner network's centralised TAT key-performance criteria, define compliance for this site; network-level TAT reports generated centrally are therefore not the denominator for this indicator. Critical Value Communication Miss: failure to communicate a critical result within the documented notification window, or incomplete documentation of the attempt; timely communication of critical results is an established post-analytical quality requirement with defined notification expectations [20]. Amended/Corrected Report: any report requiring reissue following identification of an error in the released report. Typographical Error (Manual Entry): error in manually entered demographic, request, or result information identified after release but before clinical action. Sample Retention/Storage Error (Post-Test): failure in post-testing sample storage, including temperature breaches, loss before the retention period, or storage labelling errors. LIS/Software Failure (Report Release Impact): any LIS or related software failure disrupting report release, including server crashes, software hangs, connectivity failures, and data-integrity events requiring manual intervention.

Categories were applied as mutually exclusive. Where an event could satisfy more than one definition — most commonly a report delayed by an LIS outage, which also meets the TAT breach definition — a hierarchical rule was applied: the proximate technical cause took precedence, so such events were classified as LIS/Software Failure and not additionally counted as TAT breaches. Classification was performed by the logging technologist at the point of identification and adjudicated at monthly aggregation by the laboratory head/pathologist, who resolved ambiguous entries. No report contributed to more than one category.

Incidents and affected reports

A single LIS or server failure may delay several reports simultaneously. The laboratory's error log records each such event as one incident line, annotated with the number of reports affected. Both quantities are reported here: the number of discrete incidents, which measures infrastructure reliability, and the number of affected reports, which measures patient-level

impact and forms the error count used for rate calculation. This distinction was not made explicit in the original manuscript and is reported in full below.

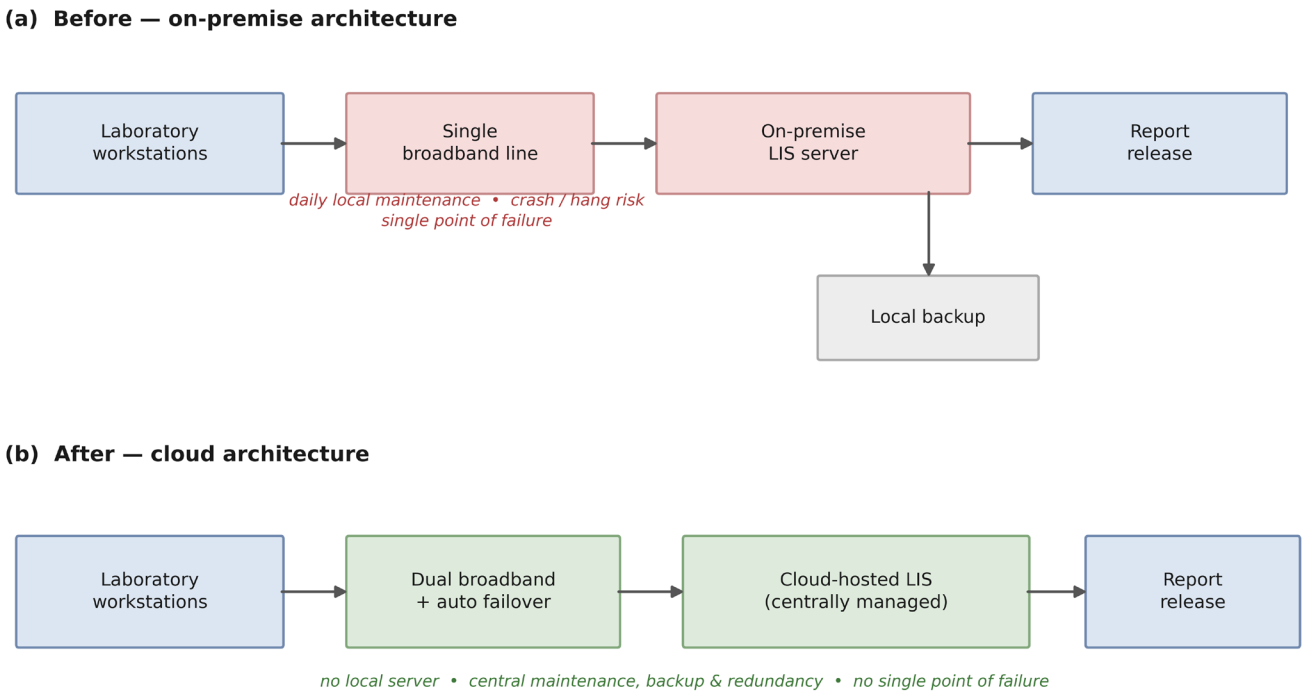
CAPA interventions (October 2025)

Interventions were implemented simultaneously across all six categories within the transitional month, and are described as a single multicomponent package. TAT Breach: a revised turnaround-time protocol with priority routing for STAT/urgent samples and escalation triggers for at-risk reports. Critical Value Communication Miss: activation of a critical-value alert in the LIS with documented notification workflow. Amended/Corrected Report: a structured sample and worksheet review workflow requiring cross-checking of patient history against current results before release. Typographical Error: role-based access control restricting review/authorisation and manual entry to authenticated personnel with audit-traceable actions. Sample Retention/Storage Error: refrigerator calibration and ongoing temperature monitoring with documented daily logs. LIS/Software Failure: the information-system architecture was redesigned through two complementary changes. First, the LIS was migrated from a local on-premise server to a centralised cloud-based LIS connected to the partner-laboratory infrastructure; the previous on-premise server required daily maintenance and was subject to crashes, software hangs and connectivity disruptions that periodically impacted report release. Second, recognising that cloud architecture creates a dependency on continuous internet uptime — under which a single line is a single point of failure — a dual-broadband configuration with automatic failover between two independent providers was installed. The migration was implemented progressively rather than as a single-day cutover. The before-and-after system architecture is contrasted in Figure 1.

Figure 2. Pareto chart of the pre-CAPA post-analytical error distribution (n = 24; April–September 2025; 19,898 reports). LIS/software failures constitute the dominant 'vital few' category, accounting for 41.7% of pre-CAPA errors; the cumulative-percentage line crosses the 80% threshold at the fourth category.

Figure 3. Run chart of monthly post-analytical error rates across the 13-month study period, expressed per 1,000 reports released using each month's actual denominator (printed on the upper axis). Absolute monthly error counts are annotated above each point, and LIS/software-failure errors are overlaid. October 2025, the implementation window, is shaded; post-analytical errors were not counted during that month. Pre- and post-CAPA medians are indicated as dotted lines.

Figure 1: Before-and-after laboratory information-system architecture.



(a) On-premise configuration: a local LIS server reached through a single broadband line, with local backup, requiring daily maintenance and presenting crash, hang and single-point-of-failure risks. (b) Cloud configuration: a centrally managed cloud-hosted LIS with central maintenance, backup and redundancy, reached through dual broadband with automatic failover, removing the local server and the single point of failure.

Because all components were deployed within the same month, their individual contributions cannot be separated. Throughout this report the intervention is therefore described as a multicomponent CAPA package, and the aggregate change is not attributed to the LIS migration alone.

Error logging, outcomes and analysis

Post-analytical errors were prospectively recorded on a standardised Monthly Post-Analytical Error Log by staff at the point of identification (date, accession identifier, category, description, immediate corrective action, logging initials), aggregated monthly by the quality-improvement coordinator, and locked at the end of each phase. The primary outcome was the post-analytical error rate per six-month period, expressed as an absolute count and as errors per million reports released (EPMR = total errors / total reports × 1,000,000).

We use the term errors per million reports released rather than Defects Per Million Opportunities. The opportunity denominator is defined explicitly as one released patient report, which is the unit at which every category in this study was logged. We acknowledge that this treats each report as an equal and independent opportunity for every error type, which is an approximation: critical-value misses can only arise where a critical value is generated, and storage errors only where a sample enters retention. Indicator-specific denominators were not captured prospectively during the study period and could not be reconstructed retrospectively with confidence; category-level rates are therefore reported on the common report denominator and should be interpreted as rates per released report rather than as true opportunity-based defect rates. This is a limitation of the present dataset.

Secondary outcomes were the category distribution pre- and post-CAPA, the Sigma level corresponding to the observed rate (standard conversion incorporating the conventional 1.5σ shift [14]), Pareto analysis of the pre-CAPA distribution [19], and the number of discrete LIS incidents. Monthly error rates, using each month's actual released-report denominator, are displayed as a run chart across the full 13-month period.

Statistical analysis. Contrary to the original submission, inferential analysis was performed. The pre- versus post-CAPA rate ratio was estimated with exact Poisson methods using the conditional binomial approach, with 95% confidence intervals; the two-sided Fisher's exact test was applied to the corresponding contingency table. Exact (Garwood) confidence intervals are reported for individual period rates. A bounding analysis was performed to quantify the maximum possible influence of excluding the transition month, by assigning hypothetical error counts to October 2025 against its actual denominator and identifying the count at which the reduction would cease to be statistically significant. Analyses were performed in Python 3 (SciPy 1.x). Given six observations either side of the intervention, no interrupted time-series model or statistical process control chart with formal control limits was fitted; the run chart is presented descriptively, and this constraint is acknowledged as a limitation.

Results

Monthly volumes and overall error reduction

Monthly released-report volumes and error counts across all 13 months are given in Table 1. Across the 40,692 reports in the two analysed periods, 24 post-analytical errors were recorded pre-CAPA (1,206 EPMR; ~4.53σ) and 4 post-CAPA (192 EPMR; ~5.05σ) — an 84.1% reduction and an improvement of approximately 0.5 Sigma levels.

The exact Poisson rate ratio was 0.159 (95% CI 0.040–0.464), corresponding to an 84.1% reduction (95% CI 53.6–96.0%); Fisher's exact test gave p = 0.00008. Exact 95% confidence intervals for the period rates were 0.76–1.79 per 1,000 reports pre-CAPA and 0.05–0.49 per 1,000 reports post-CAPA.

Table 1: Monthly released-report volumes and post-analytical error counts, April 2025 – April 2026.

Month	Phase	Reports released	Errors (n)	Rate /1000
Apr 2025	Pre-CAPA	3,100	2	0.65
May 2025	Pre-CAPA	3,024	2	0.66
Jun 2025	Pre-CAPA	3,620	7	1.93
Jul 2025	Pre-CAPA	3,231	3	0.93
Aug 2025	Pre-CAPA	3,542	4	1.13
Sep 2025	Pre-CAPA	3,381	6	1.77
Pre-CAPA total		**19,898**	**24**	**1.21**
Oct 2025	Implementation	3,419	Not counted	—
Nov 2025	Post-CAPA	3,520	1	0.28
Dec 2025	Post-CAPA	3,440	0	0.00

Month	Phase	Reports released	Errors (n)	Rate /1000
Jan 2026	Post-CAPA	3,713	1	0.27
Feb 2026	Post-CAPA	2,989	0	0.00
Mar 2026	Post-CAPA	3,910	2	0.51
Apr 2026	Post-CAPA	3,222	0	0.00
Post-CAPA total		**20,794**	**4**	**0.19**

Volumes were extracted from the laboratory information system. October 2025 was designated a priori as an implementation window; post-analytical errors were not counted during this month.

Bounding analysis for the excluded transition month
Because errors were not counted during October 2025, the effect of the exclusion was bounded rather than measured. Assigning hypothetical error counts to October against its actual denominator of 3,419 reports, and adding the month to the post-CAPA period — the most conservative treatment, attributing all transition-associated errors to the new configuration — the reduction remains statistically significant at the 5% level unless October contained 11 or more errors. The highest monthly count observed at any point during the baseline period was 7. The conclusion is therefore robust to the exclusion across the plausible range, though this analysis bounds rather than resolves the uncertainty.

FMEA risk scoring
Table 2 summarises the prospective FMEA. Four of six modes were High risk (RPN 9); TAT Breach and LIS/Software Failure scored RPN 4 and are correctly classified as Low risk under the categories defined in Methods.

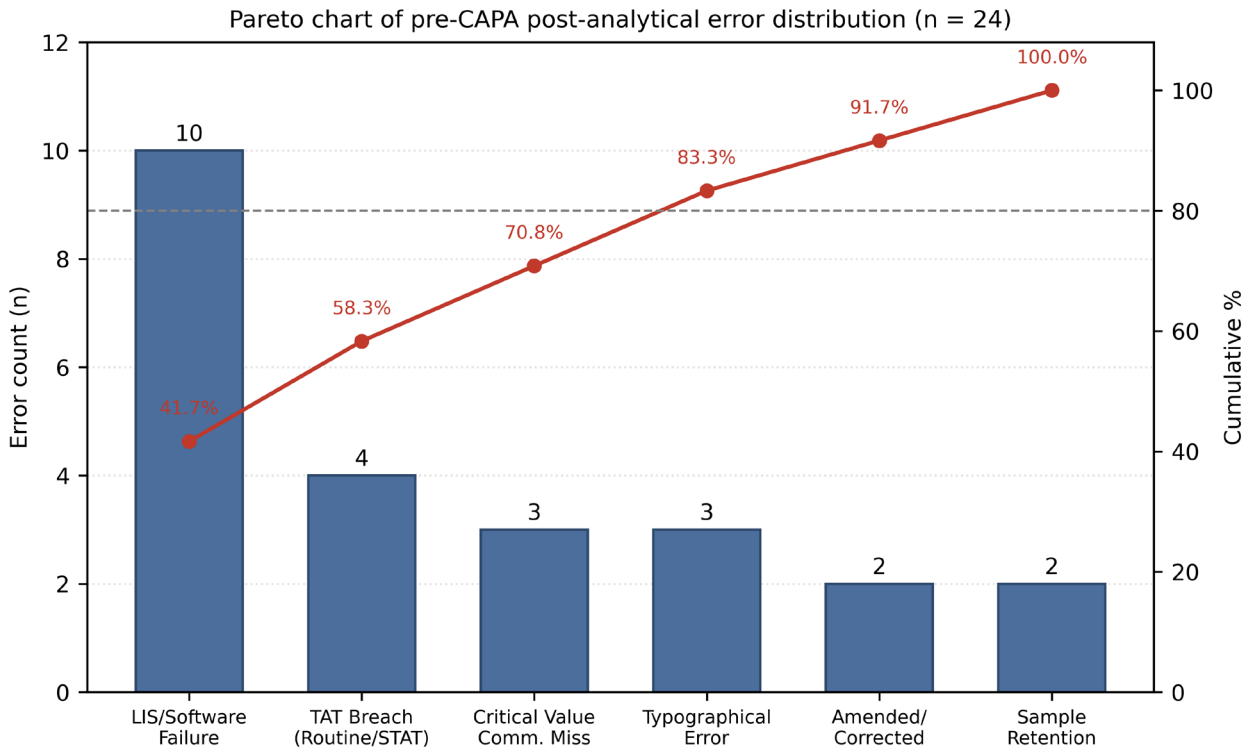
Table 2: FMEA risk scoring of post-analytical failure modes, performed March 2025 prior to study commencement.

#	Failure mode	S	P	RPN	Risk category
1	TAT Breach (Routine/STAT)	2	2	4	Low
2	Critical Value Communication Miss	3	3	9	High
3	Amended/Corrected Report	3	3	9	High
4	Typographical Error (Manual Entry)	3	3	9	High
5	Sample Retention/Storage Error	3	3	9	High
6	LIS/Software Failure (Report Release)	2	2	4	Low

S = Severity; P = Probability; RPN = Risk Priority Number (S × P). Categories: Low (<5), Medium (5–8), High (9–14), Extreme (≥15). Full anchor definitions are given in Methods and the complete matrix in Supplementary Information.

Pre-CAPA distribution and Pareto observation
LIS/software failures were the single most frequent pre-CAPA category (10/24, 41.7%), exceeding the conventional Pareto 'vital few' threshold; TAT breaches (16.7%) and typographical errors (12.5%) followed (Table 3). A Pareto chart of the pre-CAPA distribution is presented in Figure 2 .

Figure 2: Pareto chart of the pre-CAPA post-analytical error distribution (n = 24; April–September 2025; 19,898 reports).



LIS/software failures constitute the dominant 'vital few' category, accounting for 41.7% of pre-CAPA errors; the cumulative-percentage line crosses the 80% threshold at the fourth category.

Table 3: Pre-CAPA post-analytical error distribution by category (April–September 2025; 19,898 reports).

#	Error category	Count (n)	% of total	Rate /million reports
1	LIS/Software Failure	10	41.7%	503
2	TAT Breach (Routine/STAT)	4	16.7%	201
3	Critical Value Communication Miss	3	12.5%	151
4	Typographical Error	3	12.5%	151
5	Amended/Corrected Report	2	8.3%	101
6	Sample Retention/Storage Error	2	8.3%	101
	TOTAL	**24**	**100%**	**1,206**

Rates are expressed per million reports released on the common report denominator; see Methods for the limitations of this denominator.

Incident-level analysis of LIS failures

The 10 pre-CAPA LIS/software errors did not arise from 10 separate events. The error log records two discrete infrastructure incidents — one on 12 June 2025 and one on 19 September 2025, each documented as a local server failure — with five reports affected in each case. Expressed at incident level, the pre-CAPA rate was two incidents across six months (0.10 incidents per 1,000 reports); expressed at report level it was 10 affected reports (503 per million reports). No incident meeting the LIS/Software Failure definition was recorded during the six-month post-CAPA period.

This clustering explains why an infrastructure-level component contributed a large share of the absolute reduction despite a low underlying incident frequency, and it materially qualifies the apparent frequency of this failure mode. Error logs that count each affected report separately will overstate the frequency of failure modes that cluster around shared infrastructure. We report both quantities so that readers may judge infrastructure reliability and patient-level impact separately.

Post-CAPA distribution and comparison

Post-CAPA, four errors were recorded — one each in TAT Breach, Critical Value Communication Miss, Amended/Corrected Report and Sample Retention/Storage; no LIS/software failures and no typographical errors were recorded (Table 4). Every category showed a reduction, the largest absolute reduction being LIS/software failures (10 → 0).

Table 4: Pre- versus post-CAPA comparison by category and overall Sigma performance. Sigma values calculated from the observed rate using the standard Six Sigma conversion incorporating the 1.5σ shift.

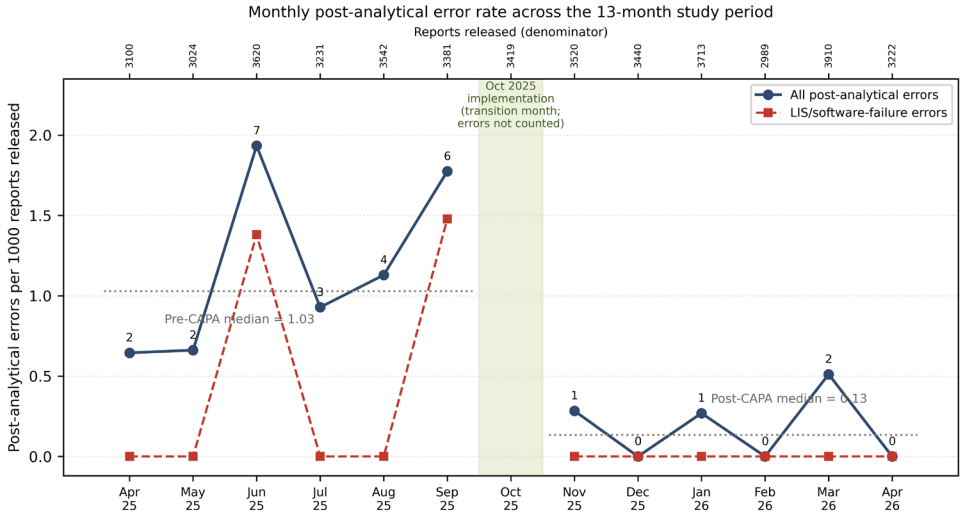
Error category	Pre-CAPA (n)	Post-CAPA (n)	Absolute Δ	% reduction
LIS/Software Failure	10	0	−10	100%
TAT Breach (Routine/STAT)	4	1	−3	75%
Critical Value Communication Miss	3	1	−2	66.7%
Typographical Error	3	0	−3	100%
Amended/Corrected Report	2	1	−1	50%
Sample Retention/Storage Error	2	1	−1	50%
TOTAL errors	**24**	**4**	**−20**	**84.1%**
Rate per million reports	**1,206**	**192**	**−1,014**	**84.1%**
Sigma performance (approx.)	**4.53σ**	**5.05σ**	**+0.52σ**	—

Category-level differences are descriptive; the small post-CAPA counts do not support category-level inference.

Monthly trajectory

The run chart (Figure 3) displays monthly error rates against each month's actual denominator. Monthly rates fell from a pre-CAPA median of 1.03 per 1,000 reports (range 0.65–1.93) to a post-CAPA median of 0.13 per 1,000 reports (range 0.00–0.51), with three post-CAPA months recording zero errors. The two highest pre-CAPA months (June and September 2025) coincided with the only months in which LIS/software failures occurred. The post-CAPA series remained at or below 0.51 per 1,000 reports across all six months. With only six observations either side of the intervention, this pattern is presented descriptively; it does not satisfy the requirements for formal run-chart or control-chart rules, and no claim of demonstrated statistical process stability is made.

Figure 3: Run chart of monthly post-analytical error rates across the 13-month study period, expressed per 1,000 reports released using each month's actual denominator (printed on the upper axis).



Absolute monthly error counts are annotated above each point, and LIS/software-failure errors are overlaid. October 2025, the implementation window, is shaded; post-analytical errors were not counted during that month. Pre- and post-CAPA medians are indicated as dotted lines.

Discussion

This 13-month before-and-after study at a stand-alone NABL-accredited laboratory found an 84.1% reduction in post-analytical errors (rate ratio 0.159, 95% CI 0.040–0.464) and an improvement in Sigma performance from ~4.53σ to ~5.05σ following a multicomponent CAPA programme. Seven distinct changes were deployed simultaneously — cloud-LIS migration, dual-broadband connectivity, revised TAT protocols, critical-value alerting, role-based access control, a structured report review workflow, and storage temperature monitoring. Because they were implemented together, the aggregate effect cannot be attributed to any single component, and we make no such attribution. The largest single categorical change was in LIS/software failures, for which no qualifying event was recorded during the post-CAPA period, but all six categories fell, which is consistent with a package effect rather than an isolated infrastructure effect.

A post-CAPA performance of approximately 5.05σ represents strong post-analytical performance relative to Sigma levels typically reported for laboratory quality indicators [15, 16]. It remains, however, materially below the six-Sigma benchmark, and we avoid describing it as approaching that benchmark.

FMEA and observed frequency

A methodologically important observation concerns the relationship between prospective risk scoring and observed error frequency. LIS/Software Failure was scored Severity 2, Probability 2 (RPN 4, Low risk). Both scores are defensible against the anchors defined in Methods: two discrete incidents across six months corresponds to approximately four events per year, matching the Probability 2 band, and the clinical consequence of an affected report was delay in release rather than an incorrect or misleading result, matching Severity 2. Yet this correctly scored low-risk mode generated 41.7% of all pre-CAPA logged errors. The explanation is that the Risk Priority Number quantifies risk per occurrence and contains no term for the number of outputs that a single occurrence affects. A laboratory allocating CAPA resource strictly by RPN would have deprioritised the largest single contributor to its logged error count, and would have been methodologically correct to do so on risk grounds. The practical implication is that prospective FMEA and retrospective error logging measure different quantities — clinical risk per event and operational error burden respectively — and neither substitutes for the other. Low clinical risk and high operational error volume can coexist in the same failure mode, particularly where that mode clusters around shared infrastructure.

Infrastructure, and the risks it introduces

Stand-alone laboratories without enterprise IT typically depend on locally administered servers maintained by laboratory staff, for which crashes, hangs and connectivity disruptions are recurrent rather than anomalous events [12, 22, 33]. Migration to a centrally managed cloud LIS transfers infrastructure maintenance to a specialised team, while dual-broadband redundancy addresses the internet-uptime dependency that cloud architecture introduces.

Cloud migration nonetheless introduces risks that laboratories considering this route must weigh explicitly [27, 34]. Data are hosted outside the laboratory's physical control, raising questions of data residency and sovereignty that differ by jurisdiction. Access control and encryption become contractual rather than local responsibilities. Vendor dependency and the practical difficulty of exporting data constitute a lock-in risk. Provider outages transfer a single point of failure from the local server room to the hosting platform, and dual-broadband redundancy does not mitigate an outage at the provider end. Disaster-recovery arrangements, breach-notification obligations and service-level commitments should be established contractually before migration, and recurring subscription costs replace one-off capital expenditure. In the present study the dual-broadband configuration used two independent commercial providers; whether their physical last-mile routes are fully disjoint was not verified, and shared physical infrastructure would reduce the effective redundancy. None of these risks materialised during the six-month post-CAPA observation period, but six months is insufficient to characterise them.

Operational adoption at other sites

Following the single-site work reported here, the same cloud-LIS migration was subsequently adopted at other laboratories within the partner network. We report this as operational adoption only. No post-analytical error outcomes were measured at those sites, no implementation timelines, technical difficulties, costs, user acceptance or inter-site variation were recorded, and no claim of demonstrated scalability or of equivalent effectiveness elsewhere is made. The original manuscript described the intervention as operationally scalable on this basis; that characterisation overstated what the observation supports and has been withdrawn.

Limitations

This study has substantial limitations, several of which constrain the strength of the conclusions.

Design and confounding. The single-centre before-and-after design cannot exclude secular trends, Hawthorne effects or co-incident operational changes. Contextual variables that could plausibly contribute to the observed change — staffing levels and turnover, test mix, number of collection centres, proportion of STAT requests, operating hours, seasonal workload variation and any concurrent quality initiatives — were not systematically captured across the two periods and cannot be compared. Their potential contribution cannot be excluded.

Excluded transition month. Post-analytical errors were not counted during October 2025. The bounding analysis reported above constrains but does not eliminate the influence of this

exclusion.

Ascertainment and surveillance bias. Errors were recorded by staff at the point of identification. The log was not subject to independent audit for unreported events, no periodic reconciliation against alternative data sources was performed, staff did not receive formal error-classification training beyond routine induction, and inter-rater reliability was not assessed. Heightened staff awareness during the post-CAPA period could plausibly have increased detection, biasing towards the null, or reduced reporting, biasing away from it. The direction of any such bias cannot be determined from the available data.

Intervention fidelity. The consistency with which each CAPA component was executed was not formally measured. Training completion, SOP adherence, critical-value notification compliance, frequency of manual overrides, broadband failover success rate and the proportion of reports receiving the mandatory cross-check were not quantified. Without fidelity measures it is not possible to state how faithfully the package was delivered.

Objective IT performance. No independent system-reliability data were collected: system uptime, number and duration of outages, latency, failed report-release attempts, backup restoration tests, automatic-failover tests, vendor service-level performance and data-integrity events were not measured. The finding regarding LIS failures rests on the absence of logged report-level errors and should be stated as such: no event meeting the LIS/Software Failure definition was recorded during the six-month post-CAPA period. This is not equivalent to independently verified system reliability, and it does not establish that such events have been permanently eliminated.

Clinical significance. All errors were weighted equally in the analysis, although their potential consequences differ substantially: a typographical error corrected before clinical action is not equivalent to a missed critical-value notification. Error severity, the distinction between near misses and errors reaching the clinician, duration of reporting delays, consequences for diagnosis or therapy, patient harm, and the number of clinicians or patients affected were not recorded and could not be reconstructed reliably. No clinical-significance stratification is therefore presented.

Unintended consequences. Potential adverse effects of the intervention were not systematically evaluated. Report authorisation time, perceived system responsiveness, staff workload, alert fatigue arising from critical-value alerting, workflow interruption from internet dependency, user satisfaction, recurring operating costs, emergence of new error types and increased dependency on central IT support were not measured.

Residual risk reassessment. A formal post-intervention re-scoring of the six failure modes was not undertaken. The FMEA reported here is the prospective assessment performed before the intervention, with the risk-category correction described in Methods; it does not represent current residual risk.

Statistical power and time series. Six observations either side of the intervention are insufficient for interrupted time-series modelling or for statistical process control with formal limits. Category-level comparisons are descriptive only.

Investigator and organisational bias. The corresponding author is the founder and managing director of the study laboratory and was directly involved in implementing the intervention, curating the data, performing the analysis and preparing the manuscript. Error classifications were adjudicated at monthly aggregation by the laboratory head/pathologist, who is a co-author but was not responsible for the infrastructure decision, and error entries were made by front-line staff rather than by the corresponding author. Beyond this, no independent verification of the error log, no audit by a quality officer external to the intervention, and no external validation of the analysis were undertaken. The potential for observer and confirmation bias is therefore mitigated only partially, and readers should weigh the findings accordingly.

Generalisability. The migration is not universally replicable for laboratories without a partner network providing cloud-LIS infrastructure, and the findings derive from a single site.

Transferability

Subject to the limitations above, the approach is expressible as a sequence other networked laboratories could consider: (i) maintain a retrospective post-analytical error log recording both incidents and affected reports, and compare it against prospective FMEA scores; (ii) recognise when interpreting prospective risk scores that the Risk Priority Number quantifies risk per occurrence and does not capture the number of reports a single event can affect, so that a correctly scored low-risk mode may still dominate the logged error count; (iii) evaluate cloud migration together with its governance, security and vendor-dependency implications rather than as a purely technical change; (iv) pair any cloud migration with independent redundant connectivity; (v) sustain process-level controls for non-infrastructure categories; and (vi) monitor the post-analytical error rate against actual monthly denominators.

Conclusion

A multicomponent CAPA programme combining cloud-LIS migration, dual-broadband redundancy and process-level controls was associated with an 84.1% reduction in post-analytical errors (rate ratio 0.159, 95% CI 0.040–0.464) and an improvement in Sigma performance from ~4.53σ to ~5.05σ over a 13-month before-and-after evaluation at a stand-alone NABL-accredited laboratory. Because all components were deployed simultaneously, the contribution of any individual component cannot be isolated. The observation that a correctly scored low-risk failure mode generated the largest share of observed errors — because the Risk Priority Number quantifies risk per occurrence and contains no term for the number of reports a single occurrence affects — supports complementing prospective risk assessment with retrospective error logging.

Multi-centre studies applying standardised post-analytical indicators, with fidelity and system-reliability measurement, are needed to substantiate these observations.

Acknowledgements
The authors thank the laboratory staff who maintained the post-analytical error log throughout the study period, and the laboratory's quality improvement coordinator for the monthly aggregation of error data.

Conflict of Interest
J.C. is the Founder and Managing Director of Vedant Diagnostic Center, the laboratory at which this study was conducted, which operates within the Agilus Diagnostics partner-laboratory network whose cloud-based laboratory information system infrastructure is described in this work; J.C. is also the Founder and Managing Director of Divyaman Biotech Private Limited. K.B. is affiliated with Vedant Diagnostic Center. These affiliations are disclosed in full. The potential for investigator bias arising from the corresponding author's dual role is addressed explicitly in the Limitations. The authors declare no other conflict of interest.

Additional Declarations

Funding
No funding was received for this work; the study was conducted on a self-funded basis at the study laboratory.

Ethics approval
This study analysed routinely collected operational quality data and involved no patient intervention, no biological sample collection and no analysis of identifiable personal data. In accordance with institutional policy and established practice for quality-improvement activities, formal institutional ethics-committee approval was not required. The work was conducted in compliance with the principles of the Declaration of Helsinki, where applicable.

Data governance
The Monthly Post-Analytical Error Log recorded the laboratory accession identifier for each entry, to permit internal verification and corrective action at the time of logging. Accession identifiers are internal laboratory numbers containing no patient name, contact detail or clinical identifier; they are, however, re-linkable to a patient record by authorised laboratory personnel within the laboratory information system, and we do not characterise them as fully anonymised. Access to the raw log was restricted to the quality-improvement coordinator, the laboratory head/pathologist and the corresponding author. The log is retained within the laboratory's NABL quality-management records for the accreditation-mandated retention period. The dataset used for the analyses reported here comprised aggregated monthly counts by category and monthly released-report denominators only; no accession identifiers or patient-level records were extracted, transferred or analysed for this study.

Data availability
The aggregated monthly error-count data, monthly denominators and FMEA scoring underlying this analysis are available from the corresponding author on reasonable request. Patient-level information was not collected and is not part of the analytical dataset.

Author contributions
J.C. and K.B. conceived and designed the study. J.C. implemented the CAPA programme and the cloud-LIS migration, curated and analysed the post-analytical error data, performed the Six Sigma, FMEA and statistical analyses, prepared Figures 1–3, and wrote the main manuscript text. K.B. contributed to the methodology, adjudicated error classifications at monthly aggregation, validated the data, supervised the study and critically revised the manuscript. Both authors reviewed and approved the final manuscript.

Use of generative AI
A large language model was used to assist with language editing, manuscript formatting and verification of statistical computations. The authors reviewed all content, take full responsibility for the work, and approved the final manuscript.

References
1. Plebani M. Errors in clinical laboratories or errors in laboratory medicine? Clin Chem Lab Med. 2006;44:750-759. doi:10.1515/CCLM.2006.123
2. Bonini P, Plebani M, Ceriotti F, Rubboli F. Errors in laboratory medicine. Clin Chem. 2002;48:691-698.
3. Carraro P, Plebani M. Errors in a stat laboratory: types and frequencies 10 years later. Clin Chem. 2007;53:1338-1342. doi:10.1373/clinchem.2007.088344
4. Plebani M. The detection and prevention of errors in laboratory medicine. Ann Clin Biochem. 2010;47:101-110. doi:10.1258/acb.2009.009222
5. Plebani M, Sciacovelli L, Aita A, Padoan A, Chiozza ML. Quality indicators to detect pre-analytical errors in laboratory testing. Clin Chim Acta. 2014;432:44-48. doi:10.1016/j.cca.2013.07.033
6. Lippi G, Blanckaert N, Bonini P, et al. Causes, consequences, detection, and prevention of identification errors in laboratory diagnostics. Clin Chem Lab Med. 2009;47:143-153. doi:10.1515/CCLM.2009.045
7. De la Salle B. Pre- and postanalytical errors in haematology. Int J Lab Hematol. 2019;41(Suppl 1):170-176. doi:10.1111/ijlh.13007
8. Sciacovelli L, Plebani M. The IFCC Working Group on laboratory errors and patient safety. Clin Chim Acta. 2009;404:79-85. doi:10.1016/j.cca.2009.03.025
9. Sciacovelli L, Padoan A, Aita A, Basso D, Plebani M.

Quality indicators in laboratory medicine: state-of-the-art, quality specifications and future strategies. Clin Chem Lab Med. 2023;61:688-695. doi:10.1515/cclm-2022-1143
10. International Organization for Standardization. ISO 15189:2022 Medical laboratories — Requirements for quality and competence. Geneva: ISO; 2022.
11. Mrazek C, Lippi G, Keppel MH, et al. Errors within the total laboratory testing process, from test selection to medical decision-making — a review of causes, consequences, surveillance and solutions. Biochem Med (Zagreb). 2020;30:020502. doi:10.11613/BM.2020.020502
12. Plebani M, Aita A, Sciacovelli L. Patient safety in laboratory medicine. In: Donaldson L, Ricciardi W, Sheridan S, Tartaglia R, eds. Textbook of Patient Safety and Clinical Risk Management. Cham: Springer; 2021:325-338. doi:10.1007/978-3-030-59403-9_24
13. National Accreditation Board for Testing and Calibration Laboratories. NABL 112 — Specific Criteria for Accreditation of Medical Laboratories. New Delhi: NABL; 2022.
14. Westgard JO, Westgard SA. Six sigma quality management system and design of risk-based statistical quality control. Clin Lab Med. 2017;37:85-96.
15. Westgard S, Bayat H, Westgard JO. Analytical sigma metrics: a review of six sigma implementation tools for medical laboratories. Biochem Med (Zagreb). 2018;28:020502. doi:10.11613/BM.2018.020502
16. Nevalainen D, Berte L, Kraft C, Leigh E, Picaso L, Morgan T. Evaluating laboratory performance on quality indicators with the six sigma scale. Arch Pathol Lab Med. 2000;124:516-519. doi:10.5858/2000-124-0516-ELPOQI
17. Chandarana J, Butani K. Risk assessment and mitigation of preanalytical errors: a quality improvement initiative at a National Accreditation Board for Testing and Calibration Laboratories-accredited diagnostic laboratory. Arch Pathol Lab Med. Published online July 21, 2026. doi:10.5858/arpa.2026-0093-OA
18. Romero A, Gómez-Salgado J, Romero-Arana A, Ruiz-Frutos C. Utilization of a healthcare failure mode and effects analysis to identify error sources in the preanalytical phase in two tertiary hospital laboratories. Biochem Med (Zagreb). 2018;28:020713. doi:10.11613/BM.2018.020713
19. Mary QA, Das S, Chaudhary N, Raju K. Analysis of the six sigma principle in pre-analytical quality for hematological specimens. Cureus. 2023;15:e42434. doi:10.7759/cureus.42434
20. Truijens K, Frans G, Vermeersch P. Critical results in laboratory medicine. Clin Chem. 2024;70:1220-1230. doi:10.1093/clinchem/hvae120
21. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. BMJ Qual Saf. 2016;25:986-992. doi:10.1136/bmjqs-2015-004411
22. Sepulveda JL, Young DS. The ideal laboratory information system. Arch Pathol Lab Med. 2013;137:1129-1140. doi:10.5858/arpa.2012-0362-RA
23. Kang HP, Sirintrapun SJ, Nestler RJ, Parwani AV. Whole slide imaging and laboratory information systems: an evolving relationship. J Pathol Inform. 2011;2:5.
24. Al-Jaroodi J, Mohamed N, Abukhousa E. Health 4.0: on the way to realizing the healthcare of the future. IEEE Access. 2020;8:211189-211210. doi:10.1109/ACCESS.2020.3038858
25. Aceto G, Persico V, Pescapé A. Industry 4.0 and health: internet of things, big data, and cloud computing for healthcare 4.0. J Ind Inf Integr. 2020;18:100129. doi:10.1016/j.jii.2020.100129
26. Coiera E. Guide to Health Informatics. 3rd ed. Boca Raton: CRC Press; 2015.
27. Kruse CS, Frederick B, Jacobson T, Monticone DK. Cybersecurity in healthcare: a systematic review of modern threats and trends. Technol Health Care. 2017;25:1-10. doi:10.3233/THC-161263
28. Sittig DF, Singh H. Electronic health records and national patient-safety goals. N Engl J Med. 2012;367:1854-1860. doi:10.1056/NEJMsb1205420
29. Bhattacherjee A, Hikmet N. Physicians' resistance toward healthcare information technology: a theoretical model and empirical test. Eur J Inf Syst. 2007;16:725-737. doi:10.1057/palgrave.ejis.3000717
30. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet. 2003;362:1225-1230. doi:10.1016/S0140-6736(03)14546-1
31. Kaplan B, Harris-Salamone KD. Health IT success and failure: recommendations from literature and an AMIA workshop. J Am Med Inform Assoc. 2009;16:291-299. doi:10.1197/jamia.M2997
32. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med. 2006;144:742-752. doi:10.7326/0003-4819-144-10-200605160-00125
33. Ulomi J, Mlay A, Kagoya HR. Laboratory information systems in resource-limited settings: implementation experiences and lessons learned. Afr J Lab Med. 2021;10:1230.
34. Chen H, Fu Z, Wang M. Data security and privacy protection issues in cloud computing. Int J Distrib Sens Netw. 2019;15:1-14.

Supplementary Information

Supplementary Table 1: Failure Mode and Effects Analysis — team composition.
The prospective FMEA was performed in March 2025, prior to commencement of the baseline observation period, by a four-member quality-improvement team:

Role	Function in the assessment
Managing Director (Quality Management)	Facilitation; review of NABL non-conformity records
Laboratory Head / Pathologist	Clinical severity judgement; adjudication of failure-mode definitions
Senior Scientific Officer	Operational frequency input from bench and reporting workflow
Quality-Improvement Coordinator	Incident-record retrieval; documentation of scoring rationale

Supplementary Table 1 FMEA team composition.

Scoring was assigned by group consensus in a single facilitated session rather than by independent rating followed by reconciliation. Disagreements were resolved by discussion until consensus was reached. No formal inter-rater reliability statistic was calculated. Detectability was not incorporated as a third dimension, because several of the failure modes assessed are by definition identified only after report release, which makes a detectability rating uninformative for this phase. These are acknowledged in the manuscript as methodological limitations.

Supplementary Table 2: Rating scale anchors as applied.

Score	Severity (S) anchor	Probability (P) anchor
1	Negligible — no clinical consequence	Rare — fewer than 1 event per year
2	Minor — administrative correction only; report delayed but not incorrect or misleading	Unlikely — 1 to 4 events per year
3	Moderate — potential to affect a clinical decision	Possible — 5 to 12 events per year
4	Major — incorrect or delayed clinical action likely	Likely — 13 to 24 events per year
5	Critical — potential for patient harm	Frequent — more than 24 events per year

Supplementary Table 2 Severity and Probability anchor definitions.

Severity (S) was scored on the clinical consequence of a single occurrence of the failure mode, that is, the consequence to the individual patient whose report was affected. It was not scored on the number of reports a single underlying event might affect.

Risk Priority Number (RPN) = S × P. Risk categories: Low (RPN < 5); Medium (RPN 5–8); High (RPN 9–14); Extreme (RPN ≥ 15).

Supplementary Table 3: Complete FMEA scoring matrix.

#	Failure mode	S	P	RPN	Risk category	Basis for the scores assigned
1	TAT Breach (Routine/ STAT)	2	2	4	Low	Severity 2: report delayed but content unaffected; result neither incorrect nor misleading. Probability 2: internal incident records indicated a small number of qualifying events per year.
2	Critical Value Communication Miss	3	3	9	High	Severity 3: failure or delay in notification can affect an urgent clinical decision. Probability 3: NABL non-conformity records showed recurrent documentation gaps.
3	Amended/Corrected Report	3	3	9	High	Severity 3: an amended report implies the original released result was wrong and may already have been acted upon. Probability 3: consistent with historical reissue frequency.
4	Typographical Error (Manual Entry)	3	3	9	High	Severity 3: manual-entry errors in demographic or result fields can mislead. Probability 3: manual entry was widespread prior to access-control changes.
5	Sample Retention/Storage Error	3	3	9	High	Severity 3: loss or degradation of a retained sample prevents repeat or add-on testing. Probability 3: temperature and retention logs showed periodic excursions.
6	LIS/Software Failure (Report Release)	2	2	4	Low	Severity 2: consequence to the affected patient is delay in report release; the released result is not altered. Probability 2: two to four qualifying infrastructure events per year were anticipated from local server incident history.

Supplementary Table 3 Complete prospective FMEA scoring matrix with rationale, as performed March 2025. Modes 1 and 6 were reported as Medium risk in the originally submitted manuscript; this was an error in the category label applied, not in the underlying S and P scores, and both are correctly classified as Low risk against the categories defined in S2. The correction is applied throughout the revised manuscript.

Supplementary Table 4: Relationship between prospective RPN and observed error frequency.

The table below compares the prospective risk classification with the observed pre-CAPA error frequency. It is presented to support the methodological point developed in the Discussion and does not constitute a residual-risk reassessment.

#	Failure mode	Prospective RPN	Risk category	Pre-CAPA errors (n)	% of pre-CAPA total	Discrete incidents
1	TAT Breach	4	Low	4	16.7%	4
2	Critical Value Communication Miss	9	High	3	12.5%	3
3	Amended/Corrected Report	9	High	2	8.3%	2
4	Typographical Error	9	High	3	12.5%	3
5	Sample Retention/Storage Error	9	High	2	8.3%	2
6	LIS/Software Failure	4	Low	10	41.7%	2

Supplementary Table 4 Prospective risk classification against observed pre-CAPA frequency. For modes 1–5 each logged error corresponds to one discrete event. For mode 6, two discrete infrastructure incidents (12 June 2025 and 19 September 2025) each affected five reports, generating ten logged errors. The Risk Priority Number quantifies risk per occurrence and contains no term for the number of reports a single occurrence affects; a mode may therefore carry correctly assessed low clinical risk while generating a high operational error count.

No formal post-intervention residual-risk reassessment was undertaken. The scoring presented here is the prospective assessment performed before the intervention and does not represent current residual risk. This is stated as a limitation in the manuscript.

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

Circulating HMGB1 as a Potential Biomarker of Inflammation in Type 2 Diabetes Mellitus and Its Complications: A Systematic Review and Meta-Analysis

Usha Kiran Pothu[1], Roshan Kumar Mahat[1*]

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

Article Info

[*]***Corresponding Author:***
Roshan Kumar Mahat, MSc (Med), PhD (Med), MBA (HHM)
Professor of Biochemistry
Teerthanker Mahaveer Medical College & Research Centre, Teerthanker Mahaveer University, Moradabad, India
E-mail: mahatroshan79@gmail.com
ORCID: 0000-0003-1202-6227

Keywords

HMGB1, type 2 diabetes mellitus, inflammation, biomarker, meta-analysis, diabetic complications

Abstract

Background: Chronic low-grade inflammation plays a pivotal role in the pathogenesis of type 2 diabetes mellitus (T2DM) and its complications. High mobility group box-1 (HMGB1), a damage-associated molecular pattern molecule, has been implicated in insulin resistance, metabolic dysregulation, and vascular injury. However, findings on circulating HMGB1 levels in T2DM are inconsistent. This systematic review and meta-analysis evaluated the association between circulating HMGB1 levels and T2DM, including its complications.

Methods: PubMed/MEDLINE, Web of Science, and Scopus were systematically searched from inception to 14 February 2026. Observational studies reporting circulating HMGB1 levels in adults with T2DM were included. A random-effects meta-analysis was performed to calculate pooled standardized mean differences (SMDs) with 95% confidence intervals (CIs). Heterogeneity was assessed using Cochran's Q, I^2, and prediction intervals. Subgroup analyses and meta-regression were conducted to explore potential sources of heterogeneity. Publication bias was evaluated using funnel plots, Egger's and Begg's tests, Doi plots with Luis Furuya-Kanamori (LFK) index, and trim-and-fill analysis.

Results: Nineteen studies were included, with sixteen studies in the primary analysis. HMGB1 levels were significantly higher in T2DM patients compared to controls (SMD = 1.39, 95% CI 0.89–1.90; $p < 0.0001$), with substantial heterogeneity ($I^2 = 96.3\%$). However, this association became non-significant after trim-and-fill adjustment. Patients with diabetic complications had significantly higher HMGB1 levels than those without (SMD = 0.52; $p < 0.0001$), particularly in microvascular complications.

Conclusion: Circulating HMGB1 levels appear to be elevated in patients with T2DM and associated complications; however, considerable heterogeneity, a prediction interval spanning the null value, and evidence of publication bias limit the certainty of this association.

Further well-designed prospective studies employing standardized methodologies are needed before HMGB1 can be considered a reliable clinical biomarker in T2DM.

Introduction

Type 2 diabetes mellitus (T2DM) is a complex metabolic disorder characterized by insulin resistance, progressive pancreatic β-cell dysfunction, and chronic low-grade inflammation. The global prevalence of T2DM continues to increase, contributing substantially to morbidity, mortality, and healthcare burden worldwide [1]. Persistent hyperglycaemia and metabolic dysregulation promote the development of microvascular complications, including diabetic nephropathy, retinopathy, and neuropathy, as well as macrovascular complications such as coronary artery disease and peripheral vascular disease [2]. Increasing evidence suggests that inflammatory mechanisms play a central role in linking metabolic disturbances with vascular injury and tissue damage in diabetes [3]. Genetic susceptibility and environmental factors further influence disease progression and complication risk [4].

High mobility group box-1 (HMGB1) is a highly conserved non-histone nuclear protein that functions as a chromatin-binding factor regulating transcription under physiological conditions [5]. During cellular stress, injury, or immune activation, HMGB1 is released extracellularly and functions as a damage-associated molecular pattern (DAMP), activating innate immune responses [6]. Extracellular HMGB1 interacts with pattern-recognition receptors such as the receptor for advanced glycation end products (RAGE) and toll-like receptors, triggering intracellular signalling cascades including nuclear factor kappa B (NF-κB) activation and inflammatory cytokine production [7]. These inflammatory pathways are implicated in metabolic dysfunction, endothelial injury, and insulin resistance [8, 9].

Emerging evidence suggests that HMGB1 is a key mediator linking hyperglycaemia-induced cellular stress with chronic inflammation and the development of diabetic complications. Experimental studies have demonstrated increased HMGB1 expression in adipose tissue, pancreatic islets, and vascular tissues under hyperglycaemic conditions, where it contributes to impaired insulin signalling, endothelial dysfunction, oxidative stress, and tissue injury [9-12]. HMGB1 has also been shown to regulate autophagy and mitochondrial function, further connecting cellular stress responses with metabolic dysregulation [13, 14]. In addition to its role in diabetes pathogenesis, elevated circulating HMGB1 levels have been associated with diabetic nephropathy through inflammatory and fibrotic mechanisms [12], retinal inflammation and angiogenesis in diabetic retinopathy [15], neuropathic changes associated with diabetic neuropathy [16], and vascular inflammation contributing to atherosclerosis and other macrovascular complications [17, 18]. These findings suggest that HMGB1 may serve not only as a mediator of chronic inflammation but also as a potential biomarker reflecting disease progression and vascular complications in T2DM.

Despite growing interest in HMGB1 as a biomarker of metabolic inflammation, findings from individual studies remain inconsistent due to variations in study design, population characteristics, and laboratory measurement methods. Therefore, systematic synthesis of available evidence is necessary to clarify the relationship between circulating HMGB1 levels and T2DM. The present systematic review and meta-analysis aimed to evaluate circulating HMGB1 levels in patients with T2DM compared with controls and to determine whether HMGB1 levels differ between T2DM patients with and without complications. Additionally, subgroup analyses were conducted to explore potential sources of heterogeneity based on geographical region, biological sample type, study design, sample size, and type of diabetic complications. Furthermore, meta-regression analyses were performed using study-level moderators, including mean age, body mass index (BMI), glycated hemoglobin (HbA1c), and diabetes duration, to investigate their potential contribution to between-study heterogeneity.

Materials and Methods

Study Design and Reporting Guidelines

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [19]. The protocol for this review was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO) under registration number CRD420261297845 (available at: https://www.crd.york.ac.uk/PROSPERO/view/CRD420261297845).

Literature Search Strategy

A comprehensive literature search was performed in the following electronic databases: PubMed/MEDLINE, Web of Science and Scopus from database inception to 14th February 2026. The search strategy combined keywords and Medical Subject Headings (MeSH) related to HMGB1 and type 2 diabetes mellitus. The following terms were used: ("Type 2 diabetes" OR "type 2 diabetes mellitus" OR "diabetes mellitus" OR "T2D" OR "T2DM") AND ("HMGB1 protein" OR "HMGB1" OR "HMG1 Protein" OR "HMG1" OR "amphoterin" OR "heparin-binding protein p30" OR "high-mobility group box 1" OR "high mobility group box protein 1"). The complete database search strategy is outlined in Supplementary Table 1. Only studies published in the English language were considered eligible for inclusion in this review. Additionally, reference lists of relevant articles and reviews were manually screened to identify potentially eligible studies.

Eligibility Criteria

Studies were included according to the PECOS framework.

Criterion	Definition
Population (P)	Adults diagnosed with type 2 diabetes mellitus.
Exposure (E)	Measurement of circulating HMGB1 levels in serum or plasma.
Comparator (C)	Healthy non-diabetic individuals or control subjects without diabetes, or normal glucose tolerant subjects.
Outcomes (O)	Primary outcome: circulating HMGB1 levels in patients with T2DM compared with controls. Secondary outcome: HMGB1 levels in T2DM with complications vs uncomplicated T2DM.
Study Design (S)	Observational studies including case–control studies and cross-sectional studies.

For studies including multiple diabetes subtypes, only data corresponding to type 2 diabetes mellitus were extracted to ensure population homogeneity.

Exclusion Criteria

Studies were excluded if they:

- Included type 1 diabetes mellitus or gestational diabetes mellitus
- Were animal or in vitro studies
- Were reviews, editorials, case reports, or conference abstracts
- Did not report quantitative HMGB1 data
- Included mixed diabetes populations without separate T2DM data
- Had no appropriate control group

Study Selection

Two reviewers independently screened the titles and abstracts of all retrieved records to identify potentially eligible studies. The full texts of relevant articles were then obtained and assessed according to the predefined inclusion criteria. Any disagreements between the two reviewers were resolved through discussion until consensus was reached.

Data Extraction

Data were independently extracted by two investigators using a standardized data extraction form. The following information was collected: first author, year of publication, country of study, study design, study population, diagnostic criteria for type 2 diabetes mellitus, HMGB1 measurement method, mean HMGB1 levels with corresponding standard deviations, type of biological sample (serum or plasma), and other relevant information. Information regarding whether HMGB1 concentrations represented crude or adjusted estimates, together with details of covariate adjustment where available, was also extracted (Supplementary Table 2). When data were presented graphically, numerical values were extracted using WebPlotDigitizer (version 5.2). Details of the studies requiring graphical data extraction are provided in Supplementary Table 3. Any discrepancies in the extracted data were resolved through discussion between the two reviewers.

Quality assessment

The methodological quality of the included studies was assessed according to study design. Case–control studies were evaluated using the Newcastle–Ottawa Scale (NOS) [20], whereas cross-sectional studies were appraised using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Analytical Cross-Sectional Studies [21]. For studies assessed using the NOS, a maximum score of nine points could be assigned across the domains of selection, comparability, and exposure/outcome assessment. Studies scoring ≤4 points were considered low quality (high risk of bias), 5–6 points moderate quality (moderate risk of bias), and ≥7 points high quality (low risk of bias). For studies evaluated using the JBI checklist, methodological quality was determined based on the proportion of affirmative (“yes”) responses. Studies with ≤49% “yes” responses were classified as having a high risk of bias (low quality), those with 50–69% as moderate risk of bias (moderate quality), and those with ≥70% as low risk of bias (high quality).

Statistical Analysis

We performed meta-analyses using a random-effects model to account for between-study variability. The standardized mean difference (SMD) with a 95% confidence interval (CI) was calculated to compare circulating HMGB1 levels between: (1) patients with type 2 diabetes mellitus (T2DM) and controls, and (2) patients with T2DM with complications and those without complications. Effect estimates were weighted using the inverse-variance method, and between-study variance was estimated using the DerSimonian–Laird approach. For studies reporting HMGB1 levels as medians with interquartile ranges or ranges, corresponding means and standard deviations were estimated using established methods [22-24]. Statistical heterogeneity was assessed using the Cochran Q test and quantified with the I^2 statistic, with $p < 0.10$ and $I^2 > 50\%$ indicating substantial heterogeneity. To characterise the dispersion of true effects, 95% prediction intervals were calculated. Subgroup analyses were conducted to explore potential sources of heterogeneity based on geographical region (China vs. other countries), sample size (<100 vs.

≥100), biological sample type (serum vs. plasma), and study design (cross-sectional vs. case–control). Differences between subgroups were assessed using the chi-square test for subgroup differences, with a significance threshold of $p < 0.05$. To further elucidate potential sources of heterogeneity, meta-regression analyses were conducted utilizing mixed-effects models, incorporating mean age, body mass index (BMI), glycated hemoglobin (HbA1c), and diabetes duration as moderators. The robustness of pooled estimates was examined through leave-one-out sensitivity analyses. Small-study effects and publication bias were assessed by visual inspection of funnel plots, alongside Egger's regression test and Begg's rank correlation test. Additionally, Doi plots were constructed, and asymmetry was quantified using the Luis Furuya-Kanamori (LFK) index [25], with values between −1 and +1 indicating symmetry, ±1 to ±2 indicating minor asymmetry, and values beyond ±2 indicating major asymmetry. To evaluate the potential impact of reporting bias, the trim-and-fill method was applied to estimate missing studies and provide bias-adjusted effect estimates. P-values <0.05 were considered statistically significant, unless otherwise specified.

Results

Characteristics of included studies

Following the predefined search strategy, a total of 1,328 records were initially identified. After removal of duplicates, 730 unique articles remained for screening. Title and abstract screening yielded 34 potentially eligible studies, which were subsequently assessed through full-text review. Ultimately, 19 studies met the inclusion criteria and were incorporated into the systematic review and meta-analysis [10, 26-43] (Figure 1). The key characteristics of the included studies are summarized in Table 1. Of the 19 included studies, 16 compared circulating HMGB1 levels between patients with T2DM and controls; these were therefore incorporated into the primary meta-analysis assessing HMGB1 levels in T2DM versus controls [10, 26, 28, 30-37, 39-43]. Five of these studies also reported HMGB1 levels in T2DM patients with and without diabetic complications [30, 31, 33, 40, 42]. Furthermore, three studies [27, 29, 38] exclusively compared T2DM patients with and without complications and did not include a healthy control group. Consequently, these three studies were excluded from the primary meta-analysis of T2DM versus controls but were included in the secondary meta-analysis comparing T2DM patients with and without complications. In total, eight studies contributed to the meta-analysis evaluating HMGB1 levels based on the presence of diabetic complications [27, 29-31, 33, 38, 40, 42].

Figure 1: PRISMA flow diagram for the study selection process.

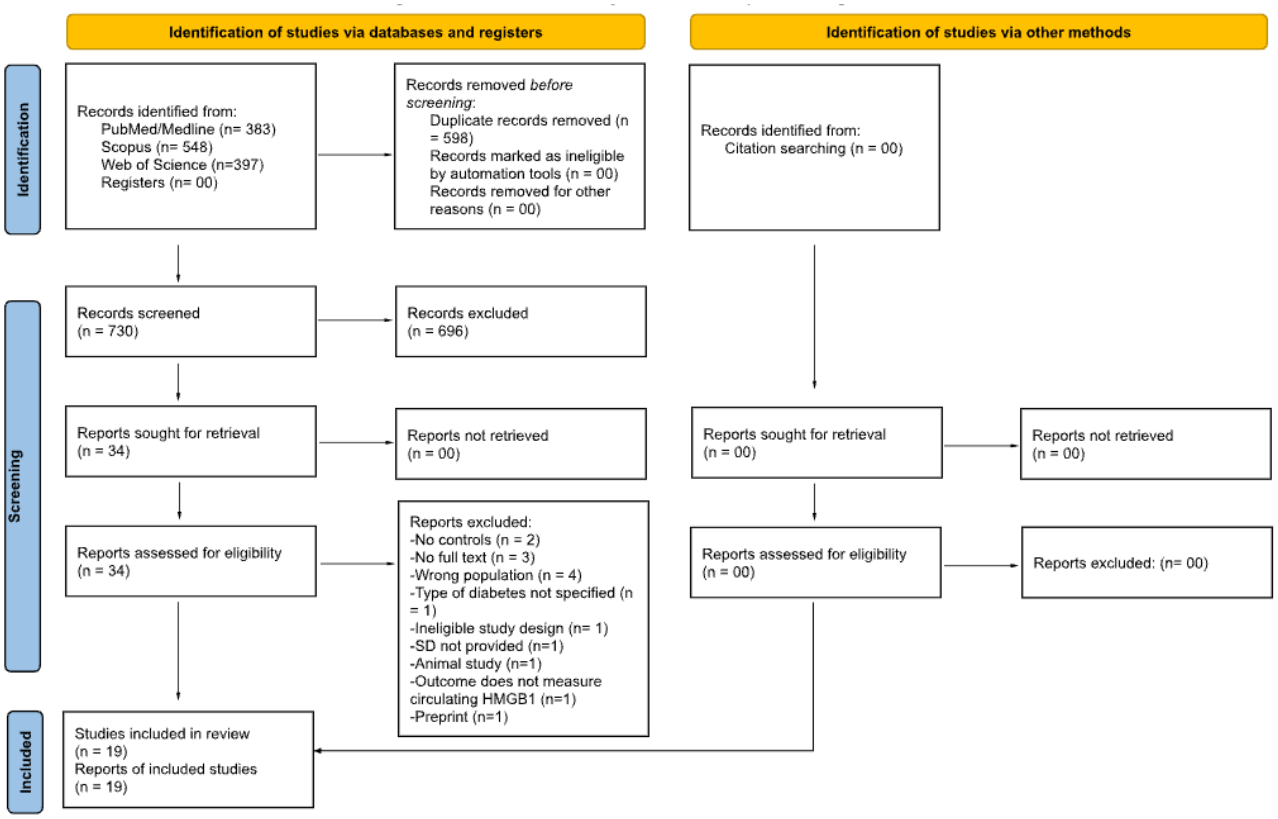


Table 1: Baseline characteristics of studies included in meta-analysis.

Author	Year	Country	Study design	Study subjects	Diagnostic criteria of type 2 DM	Sample	Main findings
Al-Hakeim HK et al. [26]	2022	Iraq	Cross-sectional study	Adults with T2DM with (n=34) and without hypertension (n=76) and healthy controls (n=51)	WHO	Serum	Serum HMGB1 levels were significantly higher in T2DM patients compared with controls and were associated with increased glucose toxicity, reduced β-cell function, and higher atherogenic indices.
Biscetti F et al. [27]	2021	Italy	Case–control study	T2DM patients with internal carotid artery stenosis (ICAS) vs T2DM patients without ICAS. Total = 873 T2DM patients (ICAS = 347, WICAS = 526)	Type 2 diabetes mellitus was defined as a fasting plasma glucose ≥126 mg/dL and/or HbA1c ≥6.5% or as medical history for the presence of diabetes plus treatment with diabetes medication.	Serum	Serum HMGB1 levels were significantly higher in T2DM patients with ICAS compared with those without ICAS, and HMGB1 was independently associated with carotid plaque vulnerability in diabetic patients.
Chen Y et al. [28]	2015	China	Case–control study	Patients with type 2 diabetes mellitus (n=50) and healthy controls (n=49)	WHO	Serum	Serum HMGB1 levels were significantly higher in T2DM patients compared with controls. HMGB1 levels were positively correlated with inflammatory cytokines IL-6 and TNF-α.
Dasu MR et al. [10]	2010	USA	Case–control study	Type 2 diabetes patients (n=23) and healthy controls (n=23)	NR	Serum	Serum HMGB1 levels were significantly higher in patients with type 2 diabetes compared with healthy controls, indicating that HMGB1 is elevated in T2DM and may contribute to the pro-inflammatory state associated with the disease.
Giovannini S. et al. [29]	2017	Italy	Case–control study	Type 2 diabetes patients with peripheral arterial disease (PAD) (n=569) vs T2DM without PAD (n=824)	HbA1c >5.8%, fasting glucose >126 mg/dL, or treatment with insulin/oral antidiabetic drugs.	Serum	Serum HMGB1 levels were significantly higher in T2DM patients with PAD compared with T2DM without PAD; HMGB1 was independently associated with PAD occurrence and severity.
Huang D et al. [30]	2020	China	Cross-sectional study	Patients with diabetes mellitus with diabetic peripheral neuropathy (DPN) (n=43), diabetes without neuropathy (n=34), incipient diabetes (n=26), and healthy controls (n=30)	WHO	Serum	Serum HMGB1 levels were significantly higher in diabetes patients compared with healthy controls and were highest in patients with diabetic peripheral neuropathy. HMGB1 levels were positively correlated with fasting glucose, postprandial glucose, HbA1c, and LDL levels.
Hui Y et al. [31]	2019	China	Cross-sectional study	Patients with type 2 diabetes mellitus with proliferative diabetic retinopathy (PDR) (n=36), non-proliferative diabetic retinopathy (NPDR) (n=43), T2DM without retinopathy (n=40), and healthy controls (n=30)	WHO	Serum	Serum HMGB1 levels were significantly elevated in patients with proliferative diabetic retinopathy compared with NPDR, T2DM without retinopathy, and healthy controls. HMGB1 levels were associated with duration of diabetes, HbA1c, and inflammatory markers.
Liu S et al. [32]	2022	China	Cross-sectional study	Newly diagnosed T2DM patients (n=40) and normal glucose tolerant (NGT) controls (n=40), further divided by obesity status	WHO	Serum	Serum HMGB1 levels were significantly higher in T2DM patients compared with normal glucose tolerant subjects. HMGB1 levels were positively correlated with BMI, fasting plasma glucose, and triglyceride levels, and negatively correlated with HDL.
Liu T et al. [33]	2024	China	Cross-sectional study	Patients with T2DM with diabetic kidney disease (DKD) (n=109), T2DM without DKD (n=87), and healthy controls (n=60)	ADA	Serum	Serum and urine HMGB1 levels were significantly higher in DKD patients compared with T2DM patients without DKD and healthy controls. Serum HMGB1 showed strong correlation with kidney injury markers (TNFR-1 and KIM-1) and renal function indicators. Elevated HMGB1 was an independent predictor of DKD progression, and levels ≥27 ng/mL were associated with a higher risk of renal function decline.
Majid H et al. (1) [34]	2025	India	Case–control study	Patients with type 2 diabetes mellitus (n=100) compared with healthy controls (n=100)	WHO and ADA diagnostic criteria	Serum	Serum HMGB1 levels were significantly higher in patients with T2DM compared with healthy controls. Increased HMGB1 levels were associated with neuroinflammatory pathways and cognitive impairment in T2DM patients.
Majid H et al. (2) [35]	2025	India	Case–control study	T2DM patients on SGLT2 inhibitors (n=57), sulfonylureas (n=54), and healthy controls (n=55)	WHO and ADA diagnostic criteria	Serum	HMGB1 and other neuroinflammatory markers were significantly elevated in T2DM patients compared to controls. SGLT2 inhibitors showed lower HMGB1 levels and better neuroinflammatory and cognitive outcomes compared to sulfonylureas.
Shen Y et al. [36]	2020	China	Cross-sectional study	Patients with proliferative diabetic retinopathy (n=32) and non-diabetic controls (n=46)	ADA	Serum	Serum and vitreous HMGB1 levels were significantly higher in diabetic retinopathy and correlated with inflammatory cytokines and angiogenic factors.
Škrha J et al. [37]	2012	Czech Republic	Case–control study	Patients with type 2 diabetes (n=66), and healthy controls (n=43)	NR	Serum	HMGB1 and other RAGE ligands were elevated in diabetic patients and associated with endothelial dysfunction markers.
Stempkowska A. et al. [38]	2020	Poland	Case–control study	T2DM patients with cardiovascular disease (n=34) vs T2DM patients without cardiovascular disease (n=60)	NR	Serum	Serum HMGB1 levels showed no significant difference between T2DM patients with cardiovascular disease and those without cardiovascular disease.
Wang H et al. [39]	2015	China	Cross-sectional study	Adults with normal glucose tolerance (n=75) and newly diagnosed T2DM (n=80)	WHO	Plasma	Plasma HMGB1 levels were significantly higher in T2DM compared with NGT and positively correlated with insulin resistance, waist-hip ratio, triglycerides, IL-6, and HOMA-IR.
Wang LJ et al. [40]	2011	China	Cross-sectional study	T2DM with HF (n=125), T2DM without HF (n=79), non-diabetic with HF (n=222), non-diabetic controls (n=220)	ADA	Serum	HMGB1 levels were significantly higher in HF patients (with and without T2DM) compared to controls.
Wei H et al. [41]	2016	China	Cross-sectional study	Subjects with normal glucose tolerance (n=62), and newly diagnosed T2DM (n=56)	ADA	Plasma	Plasma HMGB1 levels were higher in T2DM and were independently associated with insulin resistance and inflammatory markers such as IL-6.

Author	Year	Country	Study design	Study subjects	Diagnostic criteria of type 2 DM	Sample	Main findings
Yan XX et al. [42]	2009	China	Cross-sectional study	Non-diabetic without CAD (n=132), non-diabetic with CAD (n=149), T2DM without CAD (n=80), and T2DM with CAD (n=151)	Two fasting plasma glucose levels ≥7.0 mmol/L or two 2-h postprandial plasma glucose ≥11.1 mmol/L after a glucose load of 75 g or two casual glucose readings ≥11.1 mmol/L, or taking oral hypoglycemic drugs or parenteral insulin	Serum	Serum HMGB1 levels were significantly higher in patients with CAD. Among diabetic patients, HMGB1 levels were significantly higher in those with CAD compared to those without CAD (P < 0.001), indicating an association between elevated HMGB1 and diabetic coronary artery disease.
Yang X et al. [43]	2020	China	Cross-sectional study	Type 2 diabetes mellitus (n=90), and age-matched healthy controls (n=81)	WHO	Plasma	HMGB1 levels were significantly higher in T2DM patients compared with healthy controls. HMGB1 levels were associated with cognitive impairment and inflammatory processes in diabetic patients.

T2DM: Type 2 Diabetes Mellitus; HMGB1: High Mobility Group Box 1; WHO: World Health Organization; ADA: American Diabetes Association; NR: Not Reported; ICAS: Internal Carotid Artery Stenosis; WICAS: Without Internal Carotid Artery Stenosis; PAD: Peripheral Arterial Disease; DPN: Diabetic Peripheral Neuropathy; DKD: Diabetic Kidney Disease; PDR: Proliferative Diabetic Retinopathy; NPDR: Non-Proliferative Diabetic Retinopathy; CAD: Coronary Artery Disease; HF: Heart Failure; HbA1c: Glycated Hemoglobin; IL-6: Interleukin-6; TNF-α: Tumor Necrosis Factor Alpha; LDL: Low-Density Lipoprotein; HDL: High-Density Lipoprotein; BMI: Body Mass Index; HOMA-IR: Homeostatic Model Assessment of Insulin Resistance; TNFR-1: Tumor Necrosis Factor Receptor 1; KIM-1: Kidney Injury Molecule-1; NGT: Normal Glucose Tolerance. Three studies (Biscetti F et al. [27], Giovannini S et al. [29], and Stempkowska A et al. [38]) were excluded from the primary meta-analysis of T2DM versus healthy controls because they did not include a healthy control group and only compared T2DM patients with and without diabetic complications.

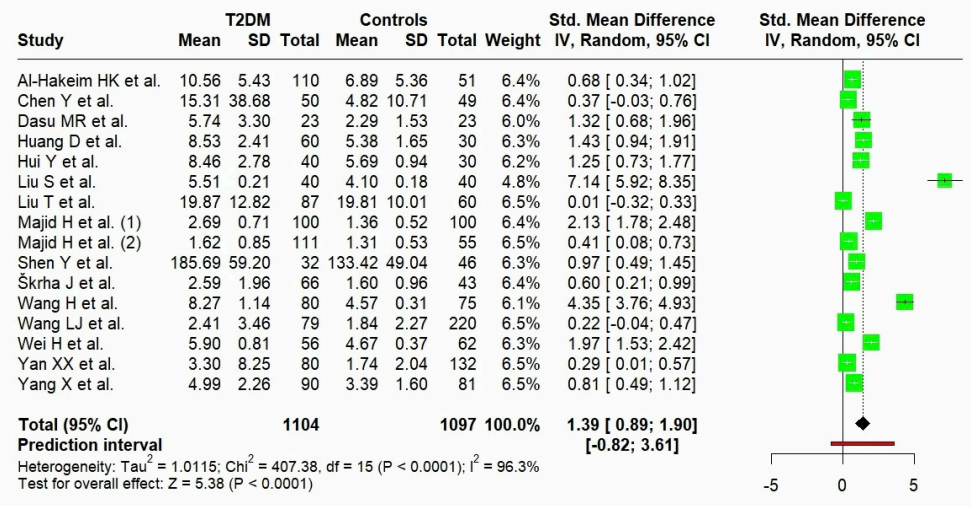
These studies were published between 2009 and 2025 and were geographically diverse, with the majority conducted in China (n = 11) [28, 30-33, 36, 39-43], followed by Italy (n = 2) [27, 29] and India (n = 2) [34, 35], and one study each from Iraq [26], the United States [10], the Czech Republic [37], and Poland [38]. In terms of study design, 11 studies were cross-sectional [26, 30-33, 36, 39-43], while 8 employed a case–control design [10, 27-29, 34, 35, 37, 38]. Regarding sample type, 16 studies measured HMGB1 in serum [10, 26-38, 40, 42] and 3 in plasma [39, 41, 43]. HMGB1 levels were assessed using enzyme-linked immunosorbent assay (ELISA) in all studies except one [28], in which the detection method was not specified. Detailed information on the ELISA kits used across studies is provided in the Supplementary Table 4. Study quality assessments are presented in Supplementary Tables 5 and 6.

HMGB1 levels between T2DM and controls

Meta-analysis

A total of 16 studies [10, 26, 28, 30-37, 39-43] including 1,104 patients with T2DM and 1,097 controls were analysed. The pooled results from the random-effects model demonstrated that circulating HMGB1 levels were significantly higher in patients with T2DM compared to controls (SMD = 1.39, 95% CI: 0.89 to 1.90, $p < 0.0001$). There was considerable between-study heterogeneity ($I^2 = 96.3\%$, $\tau^2 = 1.0115$, $p < 0.0001$), indicating substantial variability in effect sizes across studies. The prediction interval ranged from −0.82 to 3.61, suggesting that while the overall effect favors higher HMGB1 levels in T2DM, the magnitude of association may vary widely across different populations and settings. The pooled effect estimates are shown in the forest plot (Figure 2), while the distribution of statistical significance is depicted in the drapery plot (Supplementary Figure 1).

Figure 2: Forest plot showing comparison of circulating HMGB1 levels between T2DM and controls.



Subgroup analysis

When stratified by geographical region, studies conducted in China demonstrated a higher pooled effect size (SMD = 1.59, 95% CI: 0.89–2.29; I^2 = 97.1%) compared to those from other countries (SMD = 1.02, 95% CI: 0.34–1.69; I^2 = 93.5%); however, the difference between subgroups was not statistically significant (p = 0.2464). Similarly, studies with smaller sample sizes (<100 participants) showed larger effect estimates (SMD = 2.08, 95% CI: 0.77–3.39; I^2 = 96.3%) compared to larger studies (≥100 participants; SMD = 1.15, 95% CI: 0.58–1.72; I^2 = 96.6%), although subgroup differences were not statistically significant (p = 0.2025), suggesting possible small-study effects. Analysis based on biological sample type revealed higher HMGB1 levels in plasma-based studies (SMD = 2.36, 95% CI: 0.49–4.24; I^2 = 98.2%) compared to serum-based studies (SMD = 1.14, 95% CI: 0.66–1.61; I^2 = 94.9%); however, this difference was not statistically significant (p = 0.2151). Finally, subgroup analysis by study design showed that cross-sectional studies reported higher effect sizes (SMD = 1.62, 95% CI: 0.93–2.30; I^2 = 97%) compared to case–control studies (SMD = 0.96, 95% CI: 0.23–1.69; I^2 = 94.1%), although no significant subgroup difference was observed (p = 0.1967). The forest plots for these analyses are shown in Supplementary Figures 2-5.

Meta-regression

Meta-regression was conducted to ascertain whether study-level characteristics, specifically mean age, BMI, HbA1c, and diabetes duration, accounted for the observed between-study heterogeneity. None of the covariates assessed demonstrated a statistically significant influence on the pooled effect size. Mean age exhibited a non-significant inverse relationship with the effect size (β = −0.0540, 95% CI: −0.1242 to 0.0161, p = 0.1310). Similarly, BMI did not show a significant association with the pooled effect size (β = 0.0660, 95% CI: −0.1684 to 0.3004, p = 0.5811). HbA1c similarly failed to demonstrate a significant moderating effect (β = −0.4620, 95% CI: −1.0460 to 0.1221, p = 0.1211), and diabetes duration was not significantly associated with the pooled effect size (β = −0.0579, 95% CI: −0.1906 to 0.0748, p = 0.3924). Furthermore, none of the investigated moderators explained a substantial proportion of the between-study heterogeneity (R^2 = 0.0% for all models). Significant residual heterogeneity persisted across all analyses, evidenced by I^2 values ranging from 81.82% to 96.89%, thereby indicating that the variability among studies could not be adequately explained by the covariates under examination. The outcomes of the meta-regression analyses are presented in Table 2, with their corresponding graphical representations, the bubble plots, available in Supplementary Figures 6-9.

Table 2: Meta-regression analyses exploring sources of heterogeneity.

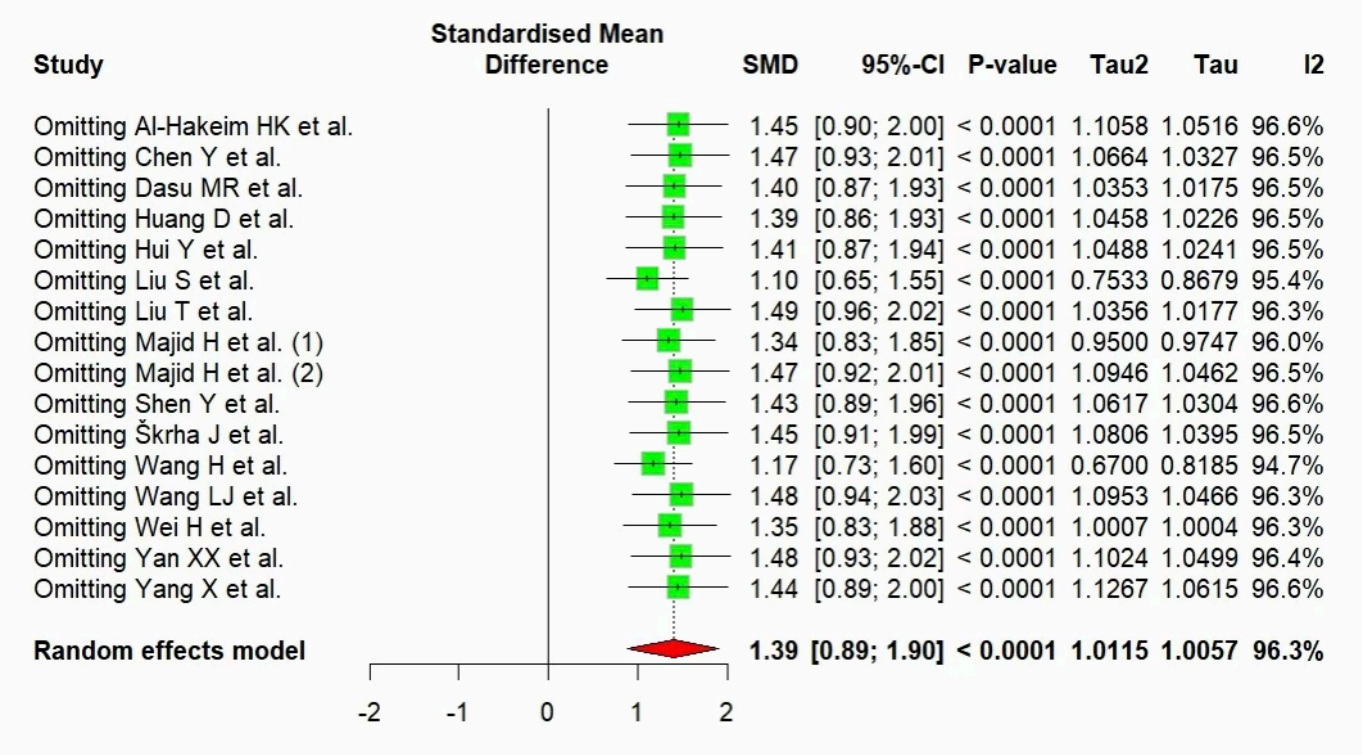
Moderator	k	β-Coefficient	SE	95% CI	QM	P-value	R^2 (%)	τ^2	I^2 (%)
Mean age	16	-0.0540	0.0358	-0.1242 to 0.0161	2.2805	0.1310	0.00	1.0718	96.42
Mean BMI	13	0.0660	0.1196	-0.1684 to 0.3004	0.3045	0.5811	0.00	1.3550	96.89
Mean HbA1c	15	-0.4620	0.2980	-1.0460 to 0.1221	2.4036	0.1211	0.00	0.7927	95.67
Diabetes duration	7	-0.0579	0.0677	-0.1906 to 0.0748	0.7314	0.3924	0.00	0.1820	81.82

β, regression coefficient; SE, standard error; CI, confidence interval; QM, test of moderators; R^2, proportion of between-study heterogeneity explained by the moderator; τ^2, residual between-study variance; I^2, residual heterogeneity; k, number of studies included in the meta-regression.

Sensitivity analysis

Leave-one-out sensitivity analysis demonstrated that the pooled effect size remained robust and statistically significant after sequential exclusion of individual studies. The pooled SMD ranged from 1.10 to 1.49, with all analyses retaining statistical significance ($p < 0.0001$). No single study disproportionately influenced the overall effect estimate (Figure 3). These findings confirm the stability and reliability of the pooled results.

Figure 3: Results of leave-one-out method in sensitivity analysis.



Publication bias

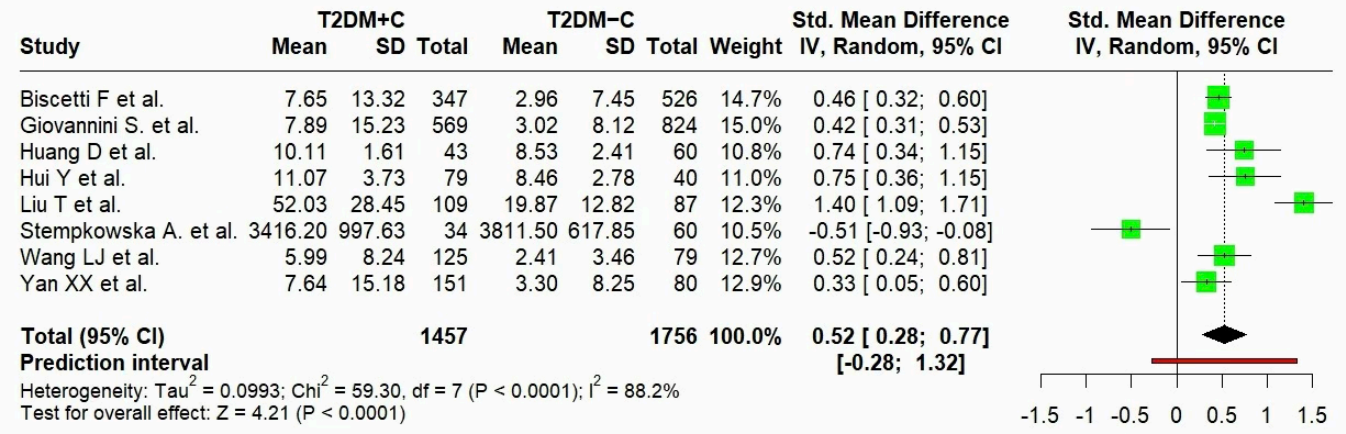
Visual inspection of the funnel plots revealed asymmetry, with an uneven distribution of studies around the pooled effect size, suggesting the presence of potential small-study effects (Supplementary Figure 10). This observation was further supported by formal statistical testing. Egger's regression test demonstrated significant funnel plot asymmetry ($t = 4.15$, $p = 0.0010$), indicating the presence of small-study effects. Similarly, Begg's rank correlation test confirmed significant asymmetry ($z = 2.97$, $p = 0.0030$), reinforcing the likelihood of publication bias. In addition, the Doi plot demonstrated major asymmetry, with an LFK index of 2.55, which is indicative of substantial publication bias (Supplementary Figure 11). To account for potential bias, a trim-and-fill analysis was performed. After adjustment, a total of 6 potentially missing studies were imputed, increasing the number of included studies to 22 (Supplementary Figure 12). The adjusted pooled effect size was attenuated and became non-significant (SMD = 0.51, 95% CI: −0.07 to 1.09; $p = 0.087$), compared to the original significant estimate. Despite adjustment, substantial heterogeneity persisted ($I^2 = 97.7\%$, $p < 0.0001$), suggesting that the observed variability is likely influenced not only by publication bias but also by clinical and methodological differences across studies.

HMGB1 levels between T2DM with complications and without complications

Meta-analysis

A total of 8 studies [27, 29-31, 33, 38, 40, 42] comprising 1,457 patients with type 2 diabetes mellitus with complications (T2DM+C) and 1,756 patients without complications (T2DM−C) were included in the meta-analysis. The pooled analysis using a random-effects model demonstrated that circulating HMGB1 levels were significantly higher in patients with T2DM with complications compared to those without complications (SMD = 0.52, 95% CI: 0.28 to 0.77, $p < 0.0001$). Substantial heterogeneity was observed across studies ($I^2 = 88.2\%$, $\tau^2 = 0.0993$). The prediction interval ranged from −0.28 to 1.32, indicating that although the overall effect favored higher HMGB1 levels in T2DM+C, the true effect may vary across different clinical settings. The results of the meta-analysis are illustrated using a forest plot (Figure 4) and a drapery plot (Supplementary Figure 13).

Figure 4: HMGB1 levels between T2DM with complications (T2DM+C) and without complications (T2DM-C).



Sensitivity analysis

Leave-one-out sensitivity analysis demonstrated that the pooled effect estimate remained statistically significant following the sequential exclusion of each individual study, indicating the robustness of the overall findings. Among the included studies, the omission of Stempkowska et al. [38] resulted in the largest change in the pooled effect estimate, increasing the standardized mean difference from 0.52 (95% CI: 0.28–0.77) to 0.63 (95% CI: 0.42–0.85), concurrently reducing heterogeneity from $I^2 = 88.2\%$ to 84.4%. Nevertheless, the direction and statistical significance of the association remained consistent, indicating that no single study exerted a substantial influence on the overall conclusions (Figure 5).

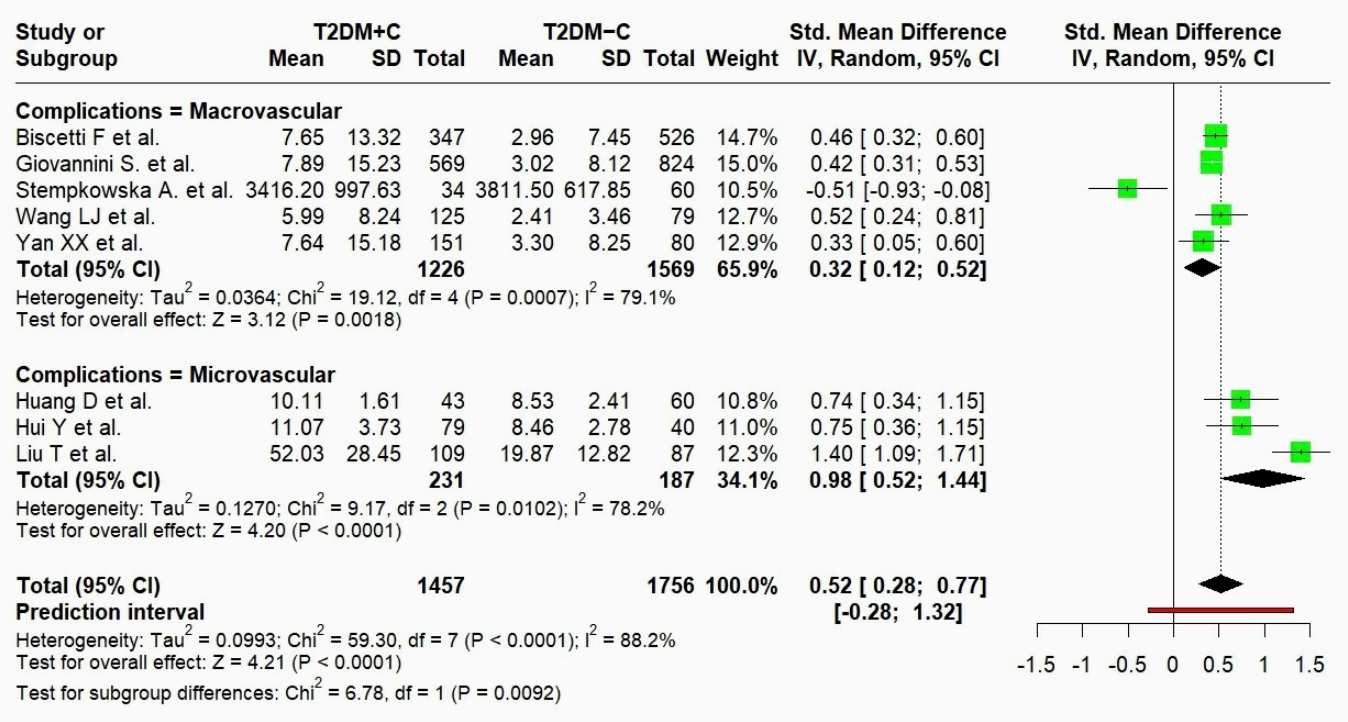
Figure 5: Sensitivity analysis illustrating the influence of individual studies on the pooled estimate for T2DM with complications.



Subgroup analysis

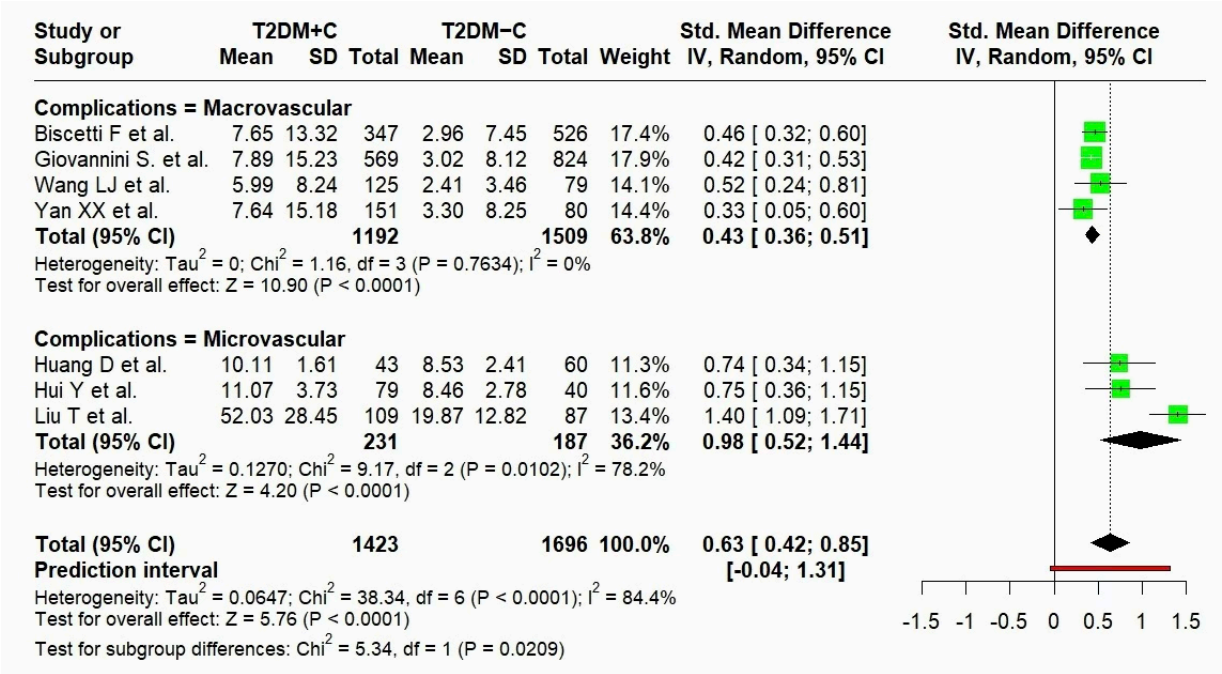
Subgroup analysis according to the type of diabetic complications demonstrated that circulating HMGB1 levels were significantly elevated in both macrovascular and microvascular complications compared with T2DM patients without complications. The pooled effect size was modest for macrovascular complications (SMD = 0.32, 95% CI: 0.12–0.52; p = 0.0018), with substantial heterogeneity (I^2 = 79.1%), whereas a considerably larger effect was observed for microvascular complications (SMD = 0.98, 95% CI: 0.52–1.44; p < 0.0001), also with substantial heterogeneity (I^2 = 78.2%). The difference between the two subgroups was statistically significant (p = 0.0092), indicating that the elevation in circulating HMGB1 levels was more pronounced in microvascular than in macrovascular complications (Figure 6).

Figure 6: Forest plot of subgroup analysis comparing circulating HMGB1 levels in patients with microvascular and macrovascular complications of T2DM versus patients without complications.



As Stempkowska et al. [38] exerted the greatest influence on the pooled estimate in the leave-one-out sensitivity analysis as shown in Figure 5, a sensitivity subgroup analysis was performed after excluding this study from the macrovascular subgroup. Exclusion of Stempkowska et al. [38] increased the pooled effect estimate for macrovascular complications from SMD = 0.32 (95% CI: 0.12–0.52) to 0.43 (95% CI: 0.36–0.51) and reduced heterogeneity from I^2 = 79.1% to 0%. The pooled estimate for microvascular complications remained unchanged, and the difference between the macrovascular and microvascular subgroups remained statistically significant (Figure 7), confirming the robustness of the subgroup findings.

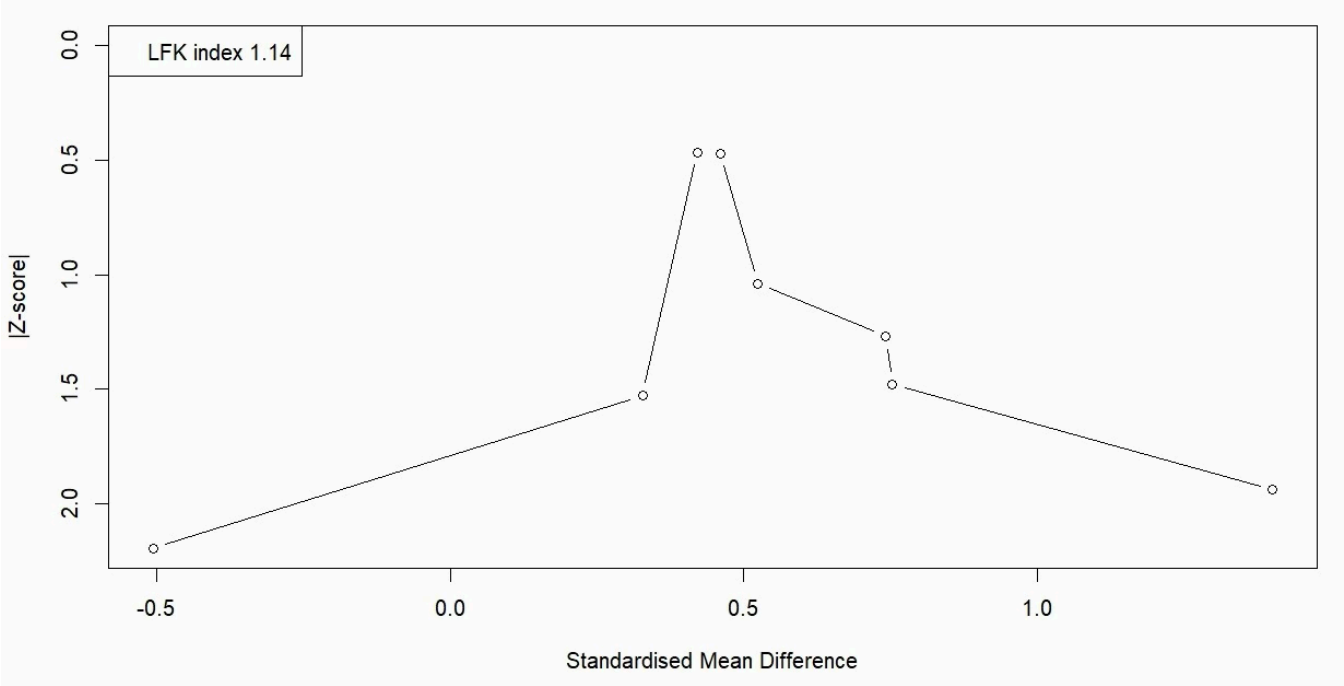
Figure 7: Forest plot of the sensitivity subgroup analysis comparing circulating HMGB1 levels in patients with microvascular and macrovascular complications of type 2 diabetes mellitus after exclusion of the influential study (Stempkowska et al. [38]).



Publication bias

Assessment of publication bias using the Doi plot demonstrated minor asymmetry, with an LFK index of 1.14, indicating low to moderate asymmetry. These findings suggest that publication bias is unlikely to have a major influence on the pooled results, although the presence of small-study effects cannot be entirely excluded (Figure 8).

Figure 8: Doi plot showing potential publication bias in T2DM with complications.



Discussion

This systematic review and meta-analysis found that circulating HMGB1 levels were higher in patients with T2DM than in non-diabetic controls in the primary analysis. These findings are consistent with the concept that chronic low-grade inflammation is a fundamental component of T2DM pathophysiology and contributes to disease progression [17,44]. The observed elevation of HMGB1 supports its potential role in metabolic dysregulation and inflammatory activation; however, the strength of this association should be interpreted cautiously because trim-and-fill analysis attenuated the pooled effect estimate to non-significance.

HMGB1 plays an important role in innate immune signalling and inflammatory amplification. Hyperglycaemia and oxidative stress stimulate active secretion of HMGB1 from immune cells and passive release from damaged tissues, promoting inflammatory signalling cascades that contribute to insulin resistance [45]. Activation of RAGE-dependent pathways further enhances cytokine production and endothelial dysfunction, linking metabolic stress with vascular injury [46]. Experimental studies have demonstrated that HMGB1-mediated activation of inflammatory pathways contributes to impaired insulin signalling and chronic metabolic inflammation [47]. Importantly, the present meta-analysis demonstrated that circulating HMGB1 levels were significantly higher in T2DM patients with complications compared with those without complications, suggesting that HMGB1 may reflect disease severity and progression. Experimental evidence indicates that HMGB1 contributes to renal inflammation and fibrosis in diabetic kidney disease, promotes inflammatory responses and pathological angiogenesis in diabetic retinopathy, and participates in neuroinflammatory pathways associated with diabetic neuropathy [12, 15, 16]. Proteomic studies further support the potential value of HMGB1 as an early biomarker of diabetic nephropathy progression [48]. In addition, HMGB1 has been implicated in macrovascular disease through its ability to promote endothelial activation, vascular smooth muscle cell proliferation, inflammatory cell recruitment, and plaque instability, thereby contributing to vascular inflammation and atherosclerosis [49, 50]. These findings suggest that HMGB1 may represent a common molecular mediator linking metabolic inflammation with vascular pathology.

A particularly noteworthy finding of the present meta-analysis was the significantly stronger association of circulating HMGB1 with microvascular complications than with macrovascular complications. This observation is biologically plausible because HMGB1 is more directly involved in the inflammatory processes underlying capillary endothelial injury. Under hyperglycaemic conditions, extracellular HMGB1 activates receptor for advanced glycation end products (RAGE)- and toll-like receptor (TLR)-mediated signalling pathways, leading to oxidative stress, endothelial dysfunction,

leukocyte recruitment, and sustained cytokine production [6, 17, 51]. These mechanisms promote increased vascular permeability, basement membrane thickening, extracellular matrix accumulation, and tissue fibrosis, which are hallmark features of diabetic nephropathy, retinopathy, and neuropathy [12, 15, 16, 48]. In contrast, although HMGB1 also contributes to macrovascular disease by promoting endothelial activation, vascular smooth muscle cell proliferation, monocyte recruitment, and plaque instability, the development of large-vessel atherosclerosis is influenced by a broader interplay of lipid metabolism, haemodynamic stress, vascular remodelling, and chronic inflammation [6, 17, 18, 49-53]. Consequently, HMGB1 may play a relatively greater role in the progression of diabetic microvascular complications than macrovascular disease, providing a biological explanation for the significantly larger pooled effect observed in the microvascular subgroup. Nevertheless, prospective mechanistic studies are required to determine whether HMGB1 has differential diagnostic or prognostic value across different diabetic vascular complications.

Substantial heterogeneity was observed across the included studies. To explore potential sources of between-study variability, both subgroup and meta-regression analyses were performed. Although studies conducted in China, studies with smaller sample sizes, plasma-based studies, and cross-sectional studies tended to report larger effect sizes than their respective counterparts, none of the subgroup differences reached statistical significance. Similarly, meta-regression analyses demonstrated that mean age, BMI, HbA1c, and diabetes duration were not significant moderators of the pooled effect estimates. These findings suggest that the observed heterogeneity could not be adequately explained by the examined study-level characteristics and may instead reflect the influence of unmeasured clinical and methodological factors. In particular, HMGB1 concentrations were measured using a variety of commercially available ELISA kits from different manufacturers, which may differ with respect to antibody specificity, calibration standards, detection limits, and analytical performance. Although the influence of assay platform could not be formally evaluated because most ELISA kits were represented by only a limited number of studies, assay-related variability should be considered when interpreting the pooled findings. Nevertheless, leave-one-out sensitivity analyses demonstrated that the pooled effect estimates remained stable following the sequential exclusion of individual studies, supporting the robustness of the results despite the presence of substantial heterogeneity.

Publication bias is another important consideration when interpreting the findings of this meta-analysis. While the primary analysis revealed significantly elevated circulating HMGB1 levels in individuals with T2DM, a trim-and-fill analysis imputed six potentially missing studies and attenuated the pooled effect estimate to a non-significant level. This finding suggests that the magnitude of the observed association may have been influenced by small-study effects and selective publication of positive results. Consequently, the current evidence may overestimate the true association between HMGB1 and T2DM. Furthermore, the wide prediction interval encompassing the null value indicates that future studies may observe smaller, null, or even opposite effects depending on the study population and clinical setting. Therefore, although HMGB1 remains a biologically plausible inflammatory biomarker involved in the pathogenesis of T2DM and its complications, the current evidence is insufficient to support its routine clinical application as a diagnostic or prognostic biomarker. Additional large-scale prospective studies using standardized assay methodologies are required to establish the true clinical utility of HMGB1.

From a clinical laboratory perspective, HMGB1 could serve as a complementary inflammatory biomarker, indicative of metabolic dysfunction and vascular injury in T2DM. Unlike established inflammatory biomarkers such as C-reactive protein (CRP), interleukin-6 (IL-6), and tumour necrosis factor-α (TNF-α), HMGB1 is directly implicated in damage-associated molecular pattern (DAMP)-mediated inflammatory signalling via RAGE and toll-like receptor pathways. This offers potential for enhanced mechanistic understanding of diabetes-related tissue injury and vascular complications. Nevertheless, current evidence is insufficient to establish its incremental diagnostic or prognostic value over existing biomarkers in routine clinical practice. Among the included studies, Liu et al. [33] identified a serum HMGB1 concentration of ≥27 ng/mL as a predictor of renal function decline in patients with T2DM, suggesting a potential prognostic role. However, this threshold lacks external validation and therefore cannot be considered a clinically applicable decision limit. Furthermore, the widespread clinical implementation of HMGB1 testing is currently constrained by the absence of standardized assay methodologies, validated reference intervals, and population-specific clinical cutoffs. Consequently, further multicentre prospective studies are warranted to standardize HMGB1 measurement, establish clinically meaningful decision thresholds, and ascertain whether HMGB1 provides incremental value over existing inflammatory biomarkers in risk stratification and disease monitoring. Emerging evidence also suggests that HMGB1 signalling pathways may represent potential therapeutic targets for reducing inflammation and preventing diabetes progression [54].

Several limitations should be acknowledged when interpreting the present meta-analysis. First, the majority of included studies were observational in design, precluding definitive causal inferences. Second, substantial heterogeneity was observed across analyses, which may limit the robustness and generalizability of the pooled estimates. This variability may partly reflect differences in assay methodologies, study populations, and clinical characteristics across studies. Third, evidence of publication bias was identified, and trim-and-fill analysis attenuated the pooled association to a non-significant

level, suggesting that the observed effect size may have been influenced by small-study effects and selective publication of positive findings. Finally, in some studies, summary statistics were not directly reported, necessitating the conversion of medians and interquartile ranges to means and standard deviations, which may introduce additional estimation bias.

Conclusion

In conclusion, circulating HMGB1 levels appear to be elevated in patients with T2DM compared with non-diabetic controls and are further increased in patients with diabetic complications compared with those without complications, particularly in microvascular complications. Nevertheless, significant between-study heterogeneity, a broad prediction interval encompassing the null value, and evidence of publication bias necessitate a cautious interpretation of these findings. Notably, trim-and-fill analysis attenuated the primary association, rendering it statistically non-significant and suggesting that the observed effect size may have been influenced by small-study effects. While HMGB1 remains a biologically plausible marker linking inflammation, metabolic dysfunction, and vascular injury, current evidence is insufficient to support its routine clinical application as a diagnostic or prognostic biomarker. Further large-scale prospective studies employing standardized measurement methodologies are required to elucidate the clinical utility of HMGB1 in T2DM.

Acknowledgements

None.

Funding

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

Declaration of Conflict of interests

The authors declare that they have no conflict of interest regarding the publication of this article.

Ethical approval

This study is a systematic review and meta-analysis. No ethical approval is required.

Credit Author Statements

Usha Kiran Pothu: Conceptualization, Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration.

Roshan Kumar Mahat: Conceptualization, Methodology, Software, Formal analysis, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration.

All authors have read and approved the final manuscript.

References

1. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. Nat Rev Immunol. 2011;11(2):98-107. https://doi.org/10.1038/nri2925
2. DeFronzo RA, Ferrannini E, Groop L, Henry RR, Herman WH, Holst JJ, et al. Type 2 diabetes mellitus. Nat Rev Dis Primers. 2015;1:15019. https://doi.org/10.1038/nrdp.2015.19
3. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-188. https://doi.org/10.1152/physrev.00045.2011
4. Cole JB, Florez JC. Genetics of diabetes mellitus and diabetes complications. Nat Rev Nephrol. 2020;16(7):377-390. https://doi.org/10.1038/s41581-020-0278-5
5. Bianchi ME. HMGB1 loves company. J Leukoc Biol. 2009;86(3):573-576. https://doi.org/10.1189/jlb.1008585
6. Andersson U, Tracey KJ. HMGB1 is a therapeutic target for sterile inflammation and infection. Annu Rev Immunol. 2011;29:139-162. https://doi.org/10.1146/annurev-immunol-030409-101323
7. Vénéreau E, Ceriotti C, Bianchi ME. DAMPs from Cell Death to New Life. Front Immunol. 2015;6:422. https://doi.org/10.3389/fimmu.2015.00422
8. Schmidt AM, Yan SD, Yan SF, Stern DM. The biology of the receptor for advanced glycation end products and its ligands. Biochim Biophys Acta. 2000;1498(2-3):99-111. https://doi.org/10.1016/s0167-4889(00)00087-2
9. Lotze MT, Tracey KJ. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. Nat Rev Immunol. 2005;5(4):331-342. https://doi.org/10.1038/nri1594
10. Dasu MR, Devaraj S, Park S, Jialal I. Increased toll-like receptor (TLR) activation and TLR ligands in recently diagnosed type 2 diabetic subjects. Diabetes Care. 2010;33(4):861-868. https://doi.org/10.2337/dc09-1799
11. Yan SF, Ramasamy R, Schmidt AM. The RAGE Axis: A fundamental mechanism signaling danger to the vulnerable vasculature. Circ Res. 2010;106:842–853.
12. Kim J, Sohn E, Kim CS, Jo K, Kim JS. The role of high-mobility group box-1 protein in the development of diabetic nephropathy. Am J Nephrol. 2011;33(6):524-529. https://doi.org/10.1159/000327992
13. Yang K, Cao F, Wang W, Tian Z, Yang L. The relationship between HMGB1 and autophagy in the pathogenesis of diabetes and its complications. Front Endocrinol (Lausanne). 2023;14:1141516. https://doi.org/10.3389/fendo.2023.1141516
14. Wahid M, Cunningham M. HMGB1 and mitochondrial dysfunction: a double-edged sword in ageing. Open Biol. 2025;15(10):240376. https://doi.org/10.1098/rsob.240376
15. Steinle JJ. Role of HMGB1 signaling in the inflammatory process in diabetic retinopathy. Cell Signal. 2020;73:109687. https://doi.org/10.1016/j.cellsig.2020.109687

16. Thakur V, Sadanandan J, Chattopadhyay M. High-Mobility Group Box 1 Protein Signaling in Painful Diabetic Neuropathy. Int J Mol Sci. 2020;21(3):881. https://doi.org/10.3390/ijms21030881
17. Biscetti F, Rando MM, Nardella E, Cecchini AL, Pecorini G, Landolfi R, et al. High Mobility Group Box-1 and Diabetes Mellitus Complications: State of the Art and Future Perspectives. Int J Mol Sci. 2019;20(24):6258. https://doi.org/10.3390/ijms20246258
18. Pengyu F, Huiyun X, Lijing G. HMGB1: a key molecule linking chronic inflammation to complications in type 2 diabetes mellitus and a target for exercise intervention. Front Endocrinol (Lausanne). 2026;17:1769497. https://doi.org/10.3389/fendo.2026.1769497
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. J Clin Epidemiol. 2021;134:178-189. https://doi.org/10.1016/j.jclinepi.2021.03.001
20. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
21. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z (Editors). Joanna Briggs Institute Reviewer's Manual. The Joanna Briggs Institute, 2017.
22. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res. 2018;27(6):1785-1805. https://doi.org/10.1177/0962280216669183
23. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135. https://doi.org/10.1186/1471-2288-14-135
24. Shi J, Luo D, Weng H, Zeng XT, Lin L, Chu H, et al. Optimally estimating the sample standard deviation from the five-number summary. Res Synth Methods. 2020;11(5):641-654. https://doi.org/10.1002/jrsm.1429
25. Furuya-Kanamori L, Barendregt JJ, Doi SA. A new improved graphical and quantitative method for detecting bias in meta-analysis. Int J Evid Based Healthc. 2018;16(4):195-203. https://doi.org/10.1097/XEB.0000000000000141
26. Al-Hakeim HK, Al-Kaabi QJ, Maes M. High mobility group box 1 and Dickkopf-related protein 1 as biomarkers of glucose toxicity, atherogenicity, and lower β cell function in patients with type 2 diabetes mellitus. Growth Factors. 2022;40(5-6):240-253. https://doi.org/10.1080/08977194.2022.2126317
27. Biscetti F, Tinelli G, Rando MM, Nardella E, Cecchini AL, Angelini F, et al. Association between carotid plaque vulnerability and high mobility group box-1 serum levels in a diabetic population. Cardiovasc Diabetol. 2021;20(1):114. https://doi.org/10.1186/s12933-021-01304-8
28. Chen Y, Qiao F, Zhao Y, Wang Y, Liu G. HMGB1 is activated in type 2 diabetes mellitus patients and in mesangial cells in response to high glucose. Int J Clin Exp Pathol. 2015;8(6):6683-6691.
29. Giovannini S, Tinelli G, Biscetti F, Straface G, Angelini F, Pitocco D, et al. Serum high mobility group box-1 and osteoprotegerin levels are associated with peripheral arterial disease and critical limb ischemia in type 2 diabetic subjects. Cardiovasc Diabetol. 2017;16(1):99. https://doi.org/10.1186/s12933-017-0581-z
30. Huang D, Zheng X, Wang S, Wu K, Xie M, Wang K. Correlations of HMGB1 and CRP with diabetic peripheral neuropathy. Rev Argent Clin Psicol. 2020;29(4):196–201. https://doi.org/10.24205/03276716.2020.821
31. Hui Y, Yin Y, Tian DH. Combination of serum HMGB-1 and serum miR-126 achieve high predictive power to detect proliferative diabetic retinopathy: a Chinese population-based study. Int J Clin Exp Med. 2019;12(5):5909–5914.
32. Liu S, Wang X, Kai Y, Tian C, Guo S, He L, et al. Clinical significance of high mobility group box 1/toll-like receptor 4 in obese diabetic patients. Endocr J. 2022;69(3):235-242. https://doi.org/10.1507/endocrj.EJ21-0381
33. Liu T, Zhao H, Wang Y, Qu P, Wang Y, Wu X, et al. Serum high mobility group box 1 as a potential biomarker for the progression of kidney disease in patients with type 2 diabetes. Front Immunol. 2024;15:1334109. https://doi.org/10.3389/fimmu.2024.1334109
34. Majid H, Islam SU, Kohli S, Nidhi. Diagnostic potential of serum biomarkers including HMGB1 and Klotho in cognitive impairment among type 2 diabetes mellitus patients: a case–control study. Beni-Suef Univ J Basic Appl Sci. 2025;14:124. https://doi.org/10.1186/s43088-025-00703-0
35. Majid H, Islam SU, Kohli S, Nidhi. Neuroinflammation and metabolic dysregulation as predictors of cognitive impairment, depression, and quality of life in type 2 diabetes mellitus patients on SGLT2 inhibitors and sulfonylureas. Inflammopharmacology. 2025;33(8):4749-4758. https://doi.org/10.1007/s10787-025-01824-9
36. Shen Y, Cao H, Chen F, Suo Y, Wang N, Xu X. A cross-sectional study of vitreous and serum high mobility group box-1 levels in proliferative diabetic retinopathy. Acta Ophthalmol. 2020;98(2):e212-e216. https://doi.org/10.1111/aos.14228
37. Skrha J Jr, Kalousová M, Svarcová J, Muravská A, Kvasnička J, Landová L, et al. Relationship of soluble RAGE and RAGE ligands HMGB1 and EN-RAGE to endothelial dysfunction in type 1 and type 2 diabetes

mellitus. Exp Clin Endocrinol Diabetes. 2012;120(5):277-281. https://doi.org/10.1055/s-0031-1283161
38. Stempkowska A, Walicka M, Franek E, Naruszewicz M, Panczyk M, Sanchak Y, et al. Hp1-1 as a Genetic Marker Regulating Inflammation and the Possibility of Developing Diabetic Complications in Patients with Type 2 Diabetes-Cohort Studies. Genes (Basel). 2020;11(11):1253. https://doi.org/10.3390/genes11111253
39. Wang H, Qu H, Deng H. Plasma HMGB-1 Levels in Subjects with Obesity and Type 2 Diabetes: A Cross-Sectional Study in China. PLoS One. 2015;10(8):e0136564. https://doi.org/10.1371/journal.pone.0136564
40. Wang LJ, Lu L, Zhang FR, Chen QJ, De Caterina R, Shen WF. Increased serum high-mobility group box-1 and cleaved receptor for advanced glycation endproducts levels and decreased endogenous secretory receptor for advanced glycation endproducts levels in diabetic and non-diabetic patients with heart failure. Eur J Heart Fail. 2011;13(4):440-449. https://doi.org/10.1093/eurjhf/hfq231
41. Wei H, Qu H, Wang H, Deng H. Plasma C1q/TNF-Related Protein-3 (CTRP-3) and High-Mobility Group Box-1 (HMGB-1) Concentrations in Subjects with Prediabetes and Type 2 Diabetes. J Diabetes Res. 2016;2016:9438760. https://doi.org/10.1155/2016/9438760
42. Yan XX, Lu L, Peng WH, Wang LJ, Zhang Q, Zhang RY, et al. Increased serum HMGB1 level is associated with coronary artery disease in nondiabetic and type 2 diabetic patients. Atherosclerosis. 2009;205(2):544-548. https://doi.org/10.1016/j.atherosclerosis.2008.12.016
43. Yang X, Chen Y, Zhang W, Zhang Z, Yang X, Wang P, et al. Association Between Inflammatory Biomarkers and Cognitive Dysfunction Analyzed by MRI in Diabetes Patients. Diabetes Metab Syndr Obes. 2020;13:4059-4065. https://doi.org/10.2147/DMSO.S271160
44. Wu H, Ma J, Wang P, Corpuz TM, Panchapakesan U, Wyburn KR, et al. HMGB1 contributes to kidney ischemia reperfusion injury. J Am Soc Nephrol. 2010;21(11):1878-1890. https://doi.org/10.1681/ASN.2009101048
45. El-Asrar AM, Nawaz MI, Kangave D, Geboes K, Ola MS, Ahmad S, et al. High-mobility group box-1 and biomarkers of inflammation in the vitreous from patients with proliferative diabetic retinopathy. Mol Vis. 2011;17:1829-1838.
46. Jia T, Zhang X, Li L, Jiang X, Wang M. Relationship Between Serum HMGB1 and RAGE Levels and Restenosis in Type 2 Diabetes Mellitus Patients Complicated With Lower Extremity Vascular Disease: A Retrospective Study. Diabetes Metab Syndr Obes. 2025;18:315-325. https://doi.org/10.2147/DMSO.S496360
47. Datta S, Rahman MA, Koka S, Boini KM. High Mobility Group Box 1 (HMGB1): Molecular Signaling and Potential Therapeutic Strategies. Cells. 2024;13(23):1946. https://doi.org/10.3390/cells13231946
48. Peng R, Zuo S, Li X, Huang Y, Chen S, Zou X, et al. Investigating HMGB1 as a potential serum biomarker for early diabetic nephropathy monitoring by quantitative proteomics. iScience. 2024;27(2):108834. https://doi.org/10.1016/j.isci.2024.108834
49. Li W, Sama AE, Wang H. Role of HMGB1 in cardiovascular diseases. Curr Opin Pharmacol. 2006;6(2):130-135. https://doi.org/10.1016/j.coph.2005.10.010
50. Rando MM, Biscetti F, Cecchini AL, Nardella E, Nicolazzi MA, Angelini F, et al. Serum high mobility group box-1 levels associated with cardiovascular events after lower extremity revascularization: a prospective study of a diabetic population. Cardiovasc Diabetol. 2022;21(1):214. https://doi.org/10.1186/s12933-022-01650-1
51. Ngcobo NN, Sibiya NH. The role of high mobility group box-1 on the development of diabetes complications: A plausible pharmacological target. Diab Vasc Dis Res. 2024;21(5):14791641241271949. https://doi.org/10.1177/14791641241271949
52. Ajoolabady A, Pratico D, Lin L, Mantzoros CS, Bahijri S, Tuomilehto J, Ren J. Inflammation in atherosclerosis: pathophysiology and mechanisms. Cell Death Dis. 2024;15(11):817. https://doi.org/10.1038/s41419-024-07166-8
53. Chawla A, Chawla R, Jaggi S. Microvasular and macrovascular complications in diabetes mellitus: Distinct or continuum? Indian J Endocrinol Metab. 2016;20(4):546-551. https://doi.org/10.4103/2230-8210.183480
54. Wang Y, Zhong J, Zhang X, Liu Z, Yang Y, Gong Q, et al. The Role of HMGB1 in the Pathogenesis of Type 2 Diabetes. J Diabetes Res. 2016;2016:2543268. https://doi.org/10.1155/2016/2543268

Supplementary Table 1: Search strategies in databases.

Databases	Search terms	No. of articles identified
PubMed/Medline	("Type 2 diabetes"[All Fields] OR "type 2 diabetes mellitus"[All Fields] OR "diabetes mellitus"[All Fields] OR "T2D"[All Fields] OR "T2DM"[All Fields]) AND ("HMGB1 protein"[All Fields] OR "HMGB1"[All Fields] OR "HMG1 Protein"[All Fields] OR "HMG1"[All Fields] OR "amphoterin"[All Fields] OR "heparin-binding protein p30"[All Fields] OR "high-mobility group box 1"[All Fields] OR "high mobility group box protein 1"[All Fields])	383
Scopus	TITLE-ABS-KEY (("Type 2 diabetes" OR "type 2 diabetes mellitus" OR "diabetes mellitus" OR "T2D" OR "T2DM") AND ("HMGB1 protein" OR "HMGB1" OR "HMG1 Protein" OR "HMG1" OR "amphoterin" OR "heparin-binding protein p30" OR "high-mobility group box 1" OR "high mobility group box protein 1"))	548
Web of Science	("Type 2 diabetes" OR "type 2 diabetes mellitus" OR "diabetes mellitus" OR "T2D" OR "T2DM") AND ("HMGB1 protein" OR "HMGB1" OR "HMG1 Protein" OR "HMG1" OR "amphoterin" OR "heparin-binding protein p30" OR "high-mobility group box 1" OR "high mobility group box protein 1")	397

Supplementary Table 2: Nature of HMGB1 estimates reported in the included studies.

Study	Values used	Covariates adjusted
Al-Hakeim HK et al. [26]	Adjusted	Age, sex, BMI, smoking
Biscetti F et al. [27]	Crude	Not applicable
Chen Y et al. [28]	Crude	Not applicable
Dasu MR et al. [10]	Crude	Not applicable
Giovannini S. et al. [29]	Crude	Not applicable
Huang D et al. [30]	Crude	Not applicable
Hui Y et al. [31]	Crude	Not applicable
Liu S et al. [32]	Crude	Not applicable
Liu T et al. [33]	Crude	Not applicable
Majid H et al. (1) [34]	Crude	Not applicable
Majid H et al. (2) [35]	Crude	Not applicable
Shen Y et al. [36]	Crude	Not applicable
Škrha J et al. [37]	Crude	Not applicable
Stempkowska A. et al. [38]	Crude	Not applicable
Wang H et al. [39]	Crude	Not applicable
Wang LJ et al. [40]	Crude	Not applicable
Wei H et al. [41]	Crude	Not applicable
Yan XX et al. [42]	Crude	Not applicable
Yang X et al. [43]	Crude	Not applicable

Covariates adjusted are reported only for studies in which the HMGB1 values used in the meta-analysis were presented as adjusted estimates.

Supplementary Table 3: Studies requiring graphical data extraction using WebPlotDigitizer (version 5.2) and the corresponding values used in the meta-analysis.

Study	Values extracted using WebPlotDigitizer	Values used in meta-analysis (Mean±SD)
Dasu MR et al. [10]	The median, minimum, and maximum HMGB1 concentrations for both the T2DM and control groups were extracted from the published figure. These values were subsequently converted to mean ± SD using established statistical methods.	Controls = 2.29±1.53 T2DM = 5.74±3.3
Hui Y et al. [31]	The mean and SD of HMGB1 concentrations were extracted for patients with proliferative diabetic retinopathy (PDR), non-proliferative diabetic retinopathy (NPDR), T2DM without retinopathy, and healthy controls. For the complications meta-analysis, the PDR and NPDR groups were combined to generate a single T2DM with complications group (T2DM+C).	Controls = 5.69±0.94 T2DM = 8.46±2.78 T2DM+C = 11.07±3.73
Liu S et al. [32]	The mean and SD of HMGB1 concentrations were extracted for the T2DM non-obese, T2DM obese, NGT non-obese, and NGT obese groups. For the primary meta-analysis, the T2DM non-obese and T2DM obese groups were combined to form a single T2DM group, while the NGT non-obese and NGT obese groups were combined to form the control group.	Controls = 4.1±0.18 T2DM = 5.51±0.21

The extracted values were independently verified before inclusion in the quantitative synthesis.

Supplementary Table 4: Details of ELISA kits used for the measurement of HMGB1.

Authors	ELISA kit used
Al-Hakeim HK et al. [26]	Elabscience, Inc. CA, USA
Biscetti F et al. [27]	Shino-Test Corporation, Tokyo
Chen Y et al. [28]	Not reported
Dasu MR et al. [10]	Shino-Test, Japan
Giovannini S. et al. [29]	Shino-Test Corporation, Tokyo
Huang D et al. [30]	Shanghai Hengyuan Biochemical Reagent Co., Ltd., China
Hui Y et al. [31]	Quantikine, XiTang, Inc., Shanghai, China
Liu S et al. [32]	Cloud-clone Corp
Liu T et al. [33]	Elabscience
Majid H et al. (1) [34]	LabReCon, New Delhi, India
Majid H et al. (2) [35]	LabReCon, New Delhi, India
Shen Y et al. [36]	Shanghai XY biotechnology Company, Ltd., China
Škrha J et al. [37]	IBL, Hamburg, Germany
Stempkowska A. et al. [38]	Human HMGB1/HMG-1 Quantikine ELISA Kit, R&D System, Abingdon, UK
Wang H et al. [39]	Uscn Life Science Inc, Wuhan, China
Wang LJ et al. [40]	Shino-Test Corporation, Tokyo, Japan
Wei H et al. [41]	Uscn Life Science Inc., Wuhan, China
Yan XX et al. [42]	Shino-Test Corporation, Tokyo
Yang X et al. [43]	R & D Systems, Minneapolis, USA

Supplementary Table 5: Quality assessment of included case-control studies by New Castle Ottawa Scale (NOS).

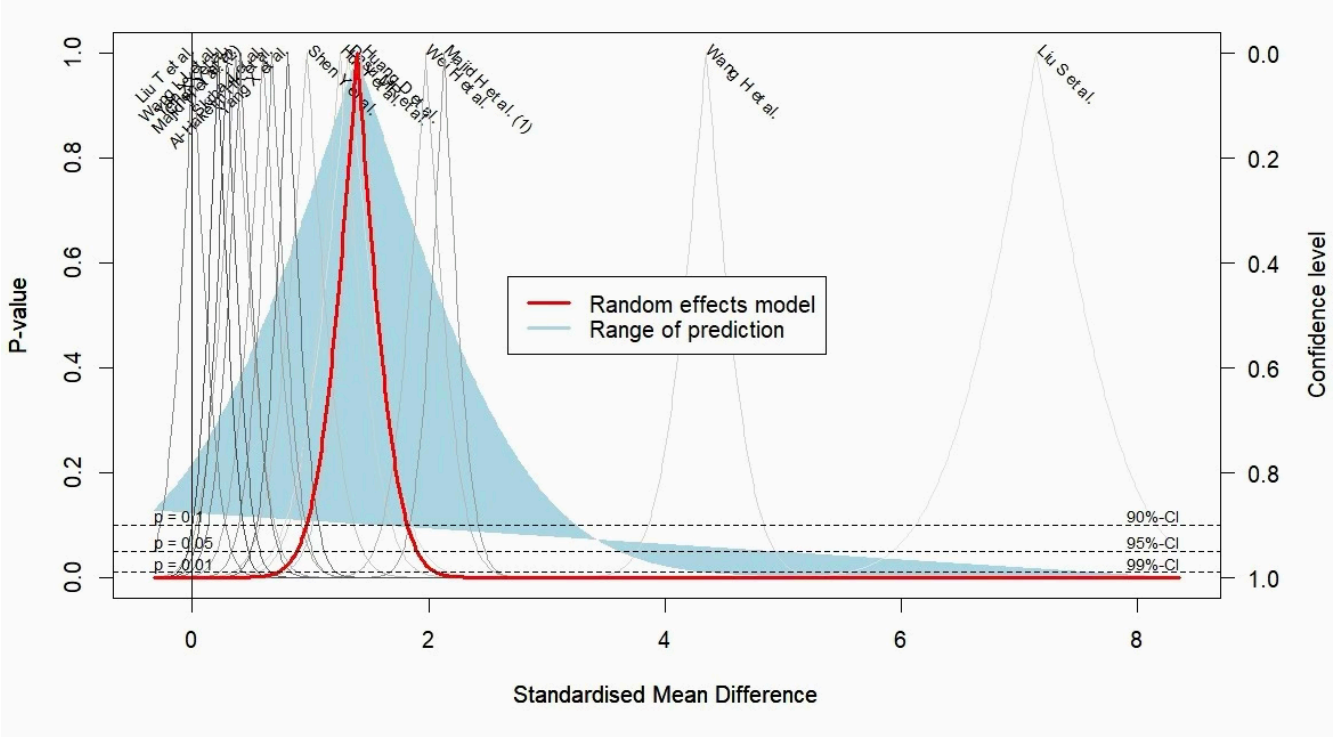
Study	Selection				Comparability		Outcome			Total Score	Risk of bias
	Adequate definition of case	Representativeness of cases	Selection of controls	Definition of controls	Age & sex	Additional factors	Ascertainment of method	Same method for both cases & controls	Non responserate		
Biscetti F et al. [27]	*	*	*	–	*	*	*	*	–	7	Low
Chen Y et al. [28]	*	*	*	*	*	–	–	*	–	6	Moderate
Dasu MR et al. [10]	*	*	–	*	*	–	*	*	–	6	Moderate
Giovannini S et al. [29]	*	*	–	*	*	*	*	*	–	7	Low
Majid H et al. [34]	*	–	–	*	*	–	*	–	*	5	Moderate
Majid H et al. [35]	*	–	–	*	*	*	*	–	–	5	Moderate
Škrha et al. [37]	*	–	–	*	*	–	*	*	–	5	Moderate
Stempkowska A. et al. [38]	*	*	–	–	–	–	*	*	–	4	High

Supplementary Table 6: Quality assessment of included cross-sectional studies by JBI Critical appraisal checklist.

Study	Were the criteria for inclusion in the sample clearly defined?	Were the study subjects and the setting described in detail?	Was the exposure measured in a valid and reliable way?	Were objective, standard criteria used for measurement of the condition?	Were confounding factors identified?	Were strategies to deal with confounding factors stated?	Were outcomes measured in a valid and reliable way?	Was appropriate statistical analysis used?	% Yes	Risk of bias
Al Hakeim HK et al. [26]	Y	Y	Y	Y	N	Y	N	N	63	Moderate
Huang D et al. [30]	N	Y	N	Y	N	N	Y	Y	50	Moderate
Hui Y et al. [31]	Y	Y	Y	Y	N	N	Y	Y	75	Low
Liu S et al. [32]	N	Y	Y	Y	N	N	N	Y	50	Moderate
Liu T et al. [33]	Y	Y	Y	Y	Y	N	Y	Y	88	Low
Shen Y et al. [36]	Y	Y	Y	Y	Y	N	Y	Y	88	Low
Wang H et al. [39]	Y	Y	Y	N	Y	Y	Y	N	75	Low
Wang L J et al. [40]	Y	Y	N	Y	Y	Y	Y	Y	88	Low
Wei H et al. [41]	Y	Y	Y	Y	Y	N	Y	Y	88	Low
Yan XX et al. [42]	Y	Y	Y	Y	N	N	Y	Y	75	Low
Yang X et al. [43]	N	Y	Y	N	N	N	Y	Y	50	Moderate

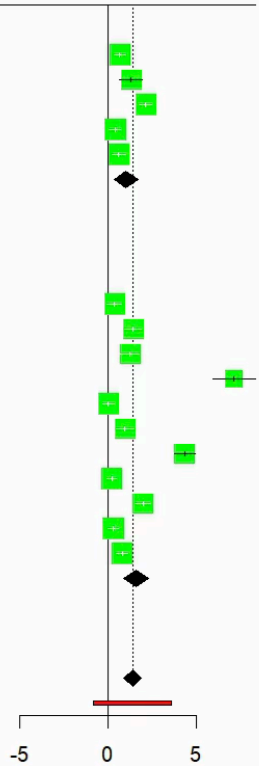
Y= Yes; N= No; Studies with ≤49% "yes" responses were classified as having a high risk of bias (low quality), those with 50–69% as moderate risk of bias (moderate quality), and those with ≥70% as low risk of bias (high quality).

Supplementary Figure 1: Drapery plot depicting the meta-analysis results based on study-specific p-value functions in T2DM.

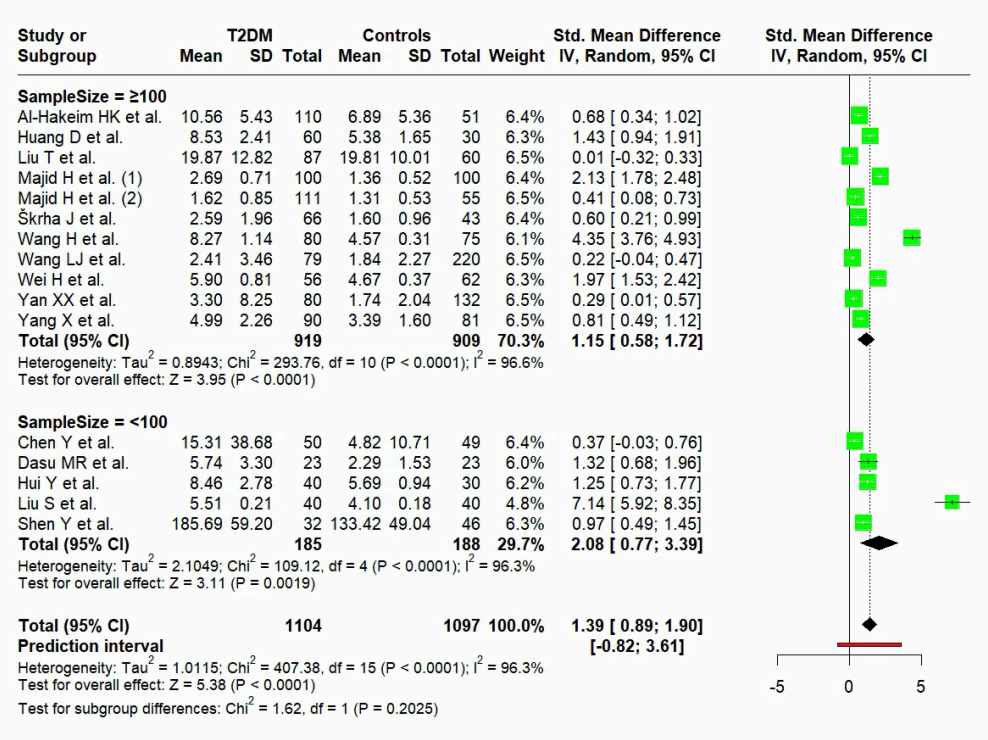


Supplementary Figure 2: Subgroup analysis by geographical region (Country).

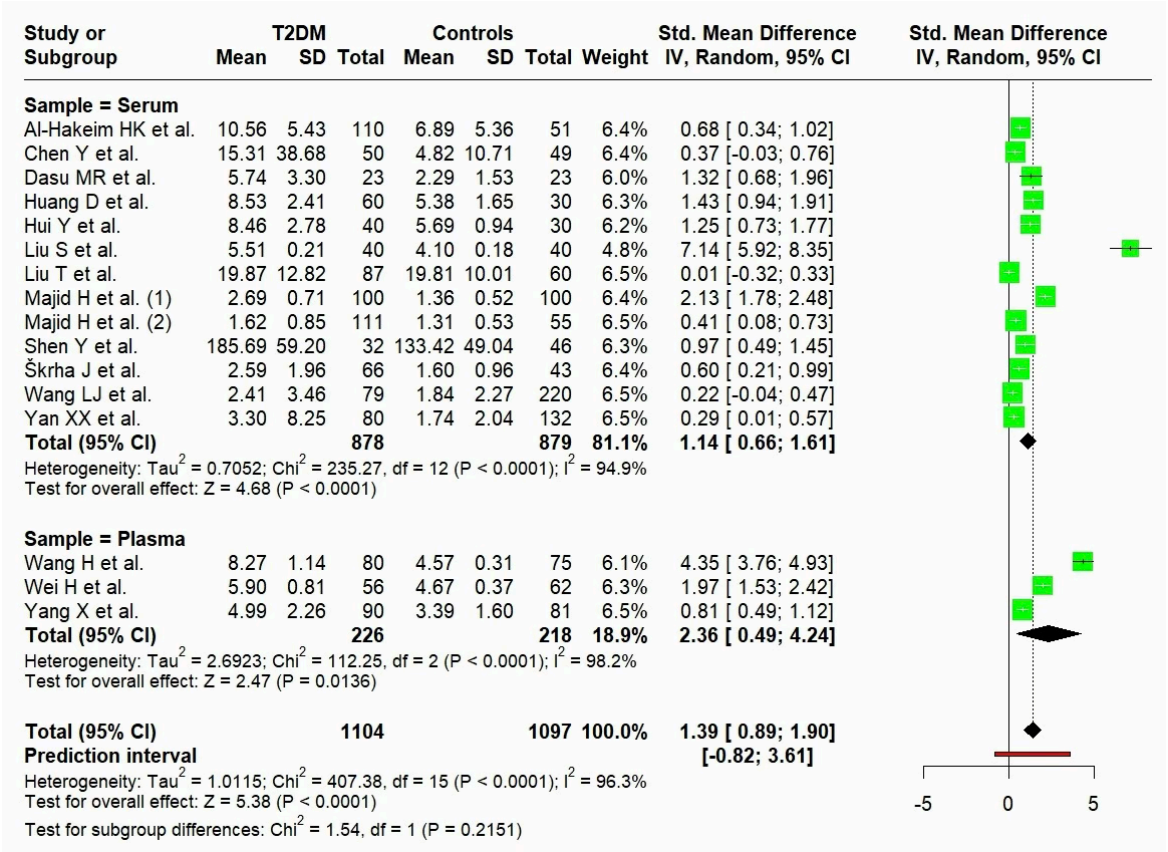
Study or Subgroup	T2DM Mean	T2DM SD	T2DM Total	Controls Mean	Controls SD	Controls Total	Weight	Std. Mean Difference IV, Random, 95% CI
Countries = Other countries								
Al-Hakeim HK et al.	10.56	5.43	110	6.89	5.36	51	6.4%	0.68 [0.34; 1.02]
Dasu MR et al.	5.74	3.30	23	2.29	1.53	23	6.0%	1.32 [0.68; 1.96]
Majid H et al. (1)	2.69	0.71	100	1.36	0.52	100	6.4%	2.13 [1.78; 2.48]
Majid H et al. (2)	1.62	0.85	111	1.31	0.53	55	6.5%	0.41 [0.08; 0.73]
Škrha J et al.	2.59	1.96	66	1.60	0.96	43	6.4%	0.60 [0.21; 0.99]
Total (95% CI)			**410**			**272**	**31.7%**	**1.02 [0.34; 1.69]**
Heterogeneity: Tau^2 = 0.5489; Chi^2 = 61.5, df = 4 (P < 0.0001); I^2 = 93.5%								
Test for overall effect: Z = 2.95 (P = 0.0031)								
Countries = China								
Chen Y et al.	15.31	38.68	50	4.82	10.71	49	6.4%	0.37 [-0.03; 0.76]
Huang D et al.	8.53	2.41	60	5.38	1.65	30	6.3%	1.43 [0.94; 1.91]
Hui Y et al.	8.46	2.78	40	5.69	0.94	30	6.2%	1.25 [0.73; 1.77]
Liu S et al.	5.51	0.21	40	4.10	0.18	40	4.8%	7.14 [5.92; 8.35]
Liu T et al.	19.87	12.82	87	19.81	10.01	60	6.5%	0.01 [-0.32; 0.33]
Shen Y et al.	185.69	59.20	32	133.42	49.04	46	6.3%	0.97 [0.49; 1.45]
Wang H et al.	8.27	1.14	80	4.57	0.31	75	6.1%	4.35 [3.76; 4.93]
Wang LJ et al.	2.41	3.46	79	1.84	2.27	220	6.5%	0.22 [-0.04; 0.47]
Wei H et al.	5.90	0.81	56	4.67	0.37	62	6.3%	1.97 [1.53; 2.42]
Yan XX et al.	3.30	8.25	80	1.74	2.04	132	6.5%	0.29 [0.01; 0.57]
Yang X et al.	4.99	2.26	90	3.39	1.60	81	6.5%	0.81 [0.49; 1.12]
Total (95% CI)			**694**			**825**	**68.3%**	**1.59 [0.89; 2.29]**
Heterogeneity: Tau^2 = 1.3288; Chi^2 = 343.27, df = 10 (P < 0.0001); I^2 = 97.1%								
Test for overall effect: Z = 4.47 (P < 0.0001)								
Total (95% CI)			**1104**			**1097**	**100.0%**	**1.39 [0.89; 1.90]**
Prediction interval								**[-0.82; 3.61]**
Heterogeneity: Tau^2 = 1.0115; Chi^2 = 407.38, df = 15 (P < 0.0001); I^2 = 96.3%								
Test for overall effect: Z = 5.38 (P < 0.0001)								
Test for subgroup differences: Chi^2 = 1.34, df = 1 (P = 0.2464)								



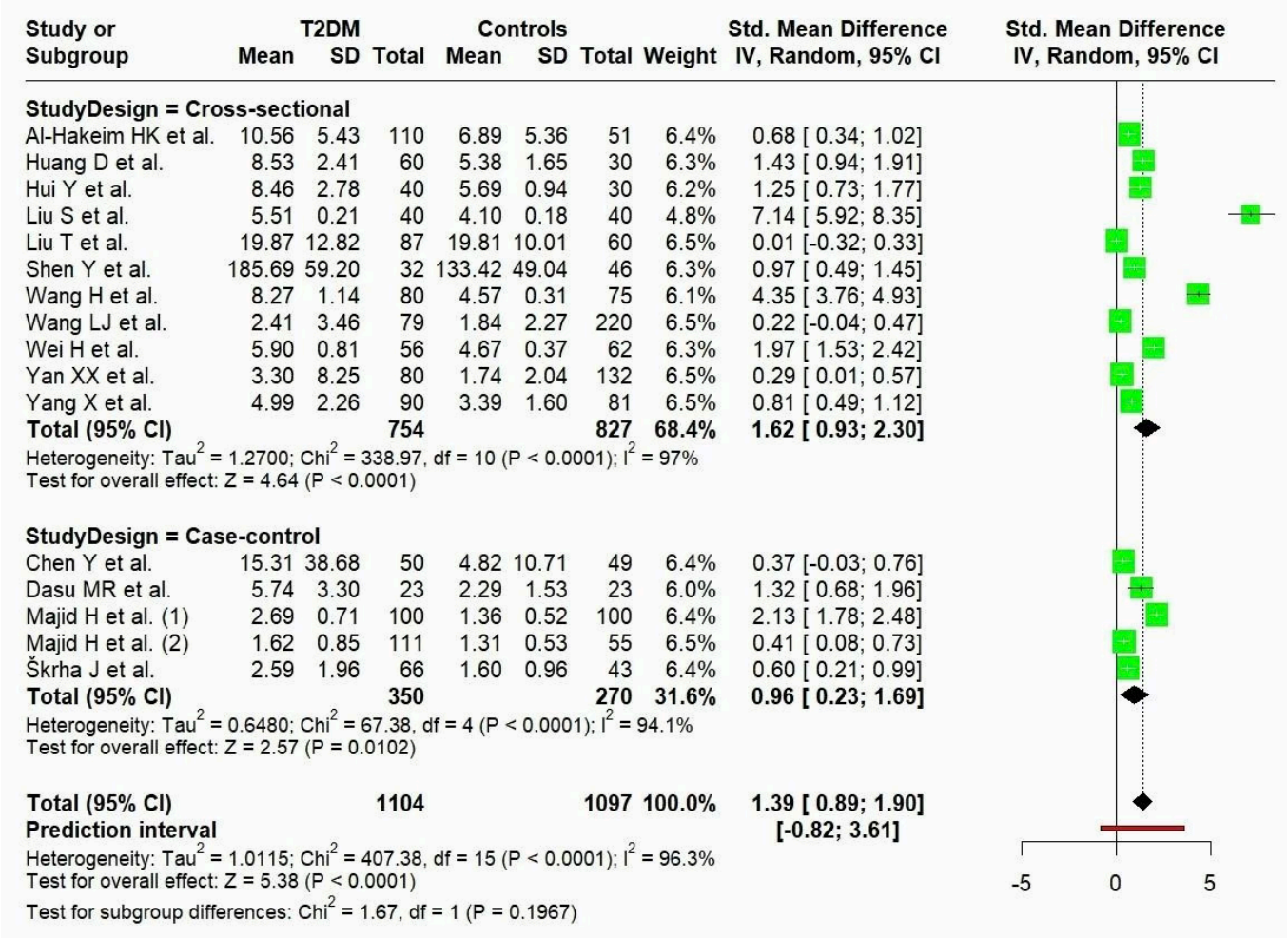
Supplementary Figure 3: Subgroup analysis by sample size.



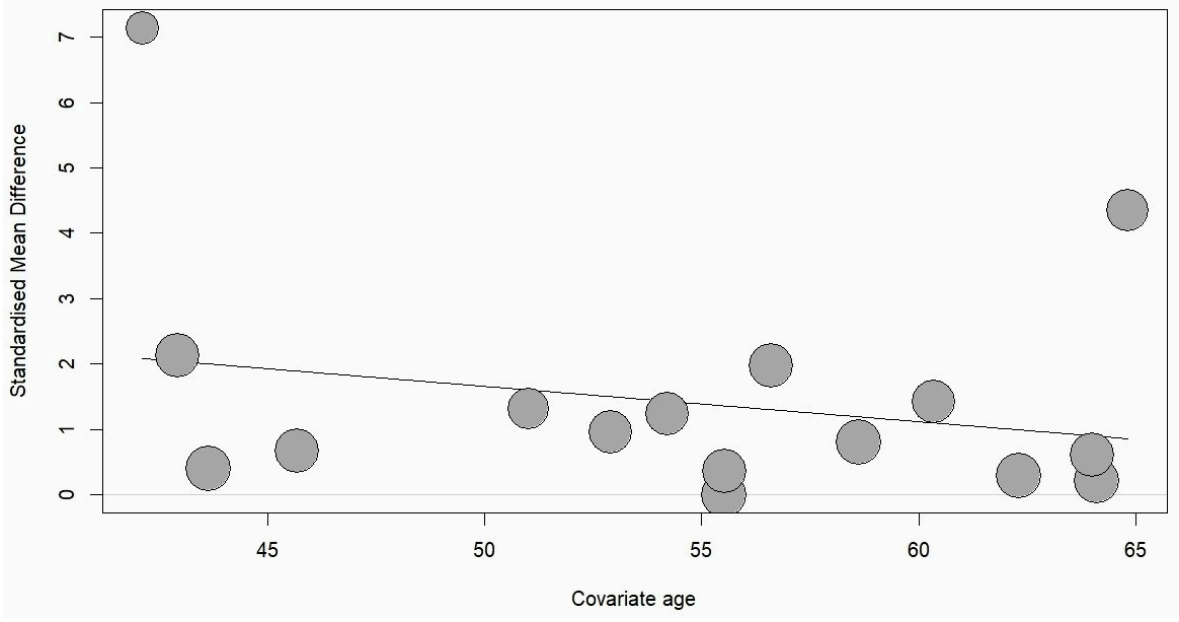
Supplementary Figure 4: Subgroup analysis by biological sample type.



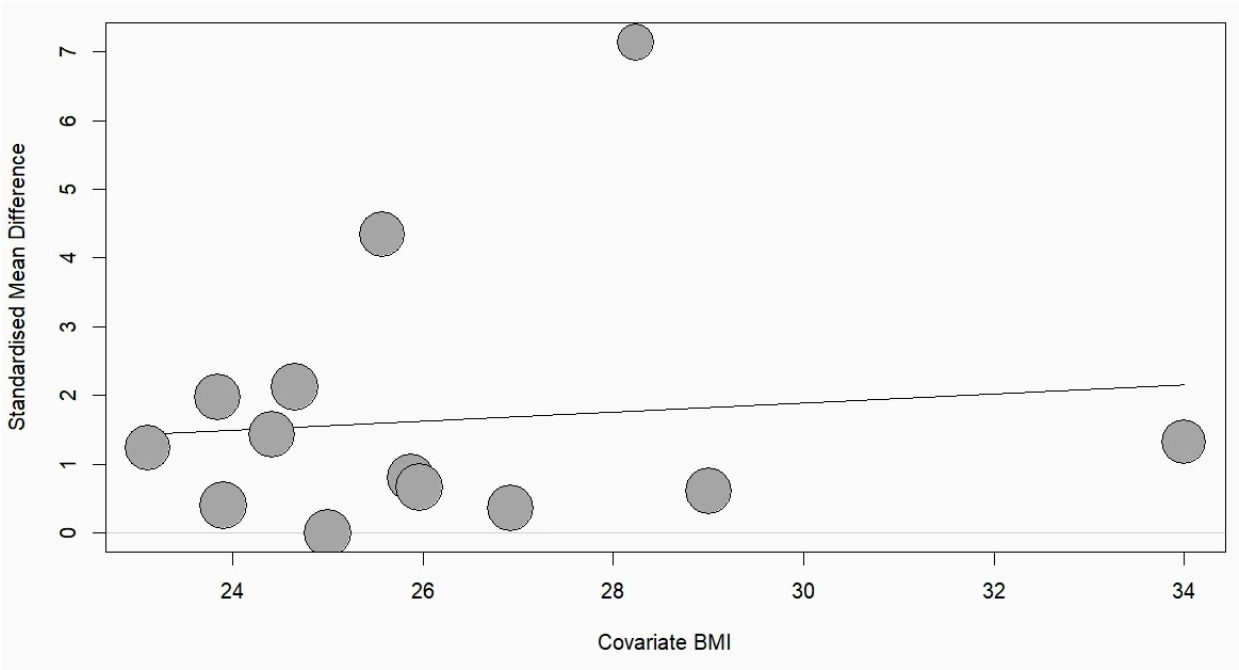
Supplementary Figure 5: Subgroup analysis by study design.



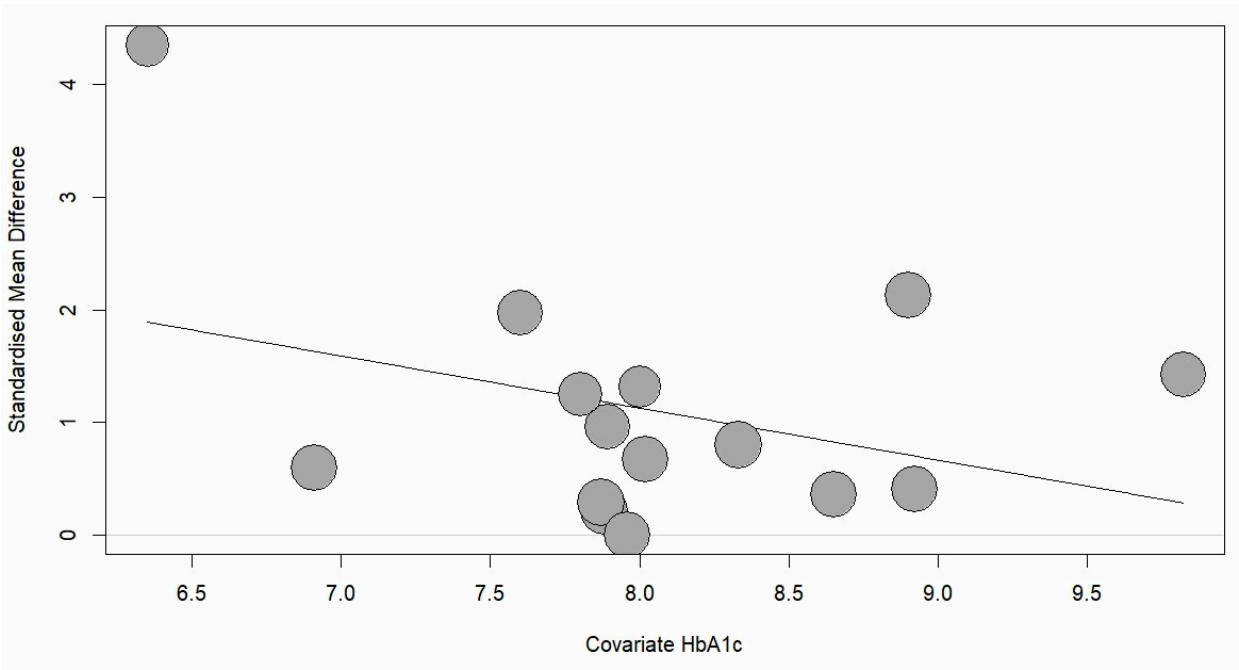
Supplementary Figure 6: Bubble plot of the meta-regression analysis evaluating the association between mean age and the standardized mean difference (SMD) of circulating HMGB1 levels in patients with T2DM.



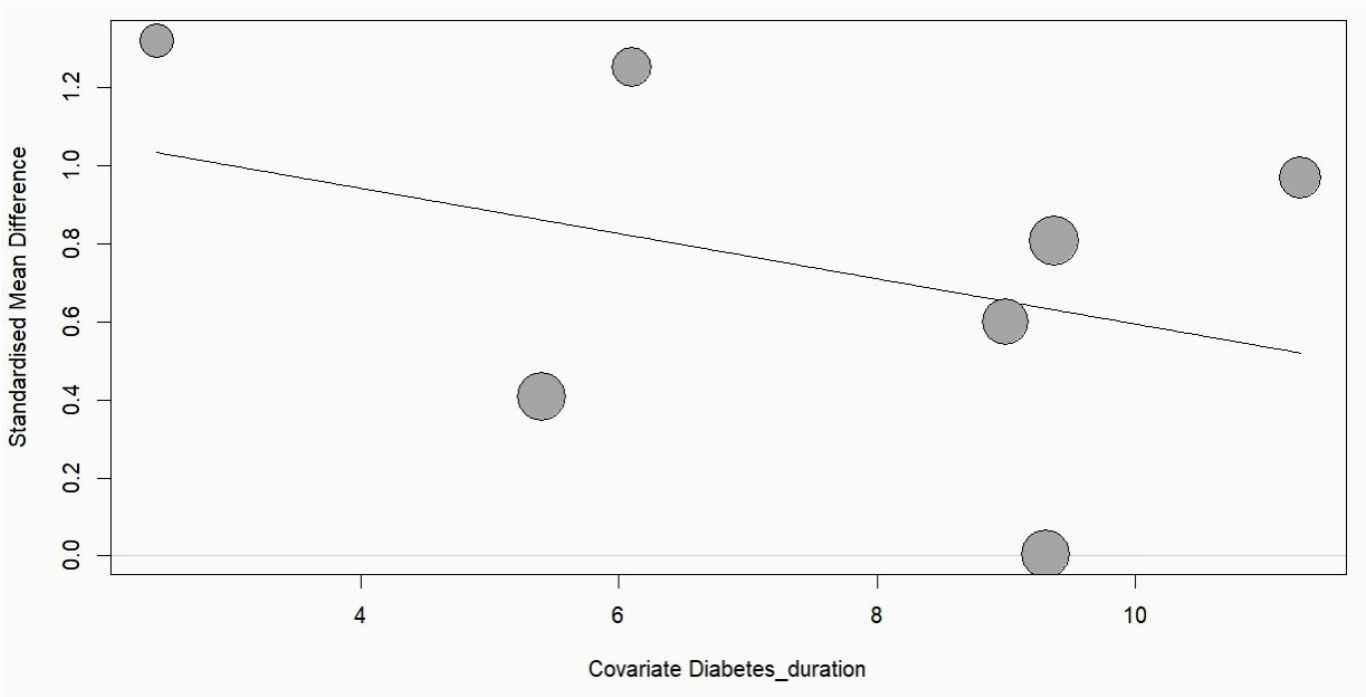
Supplementary Figure 7: Bubble plot of the meta-regression analysis evaluating the association between BMI and the standardized mean difference (SMD) of circulating HMGB1 levels in patients with T2DM.



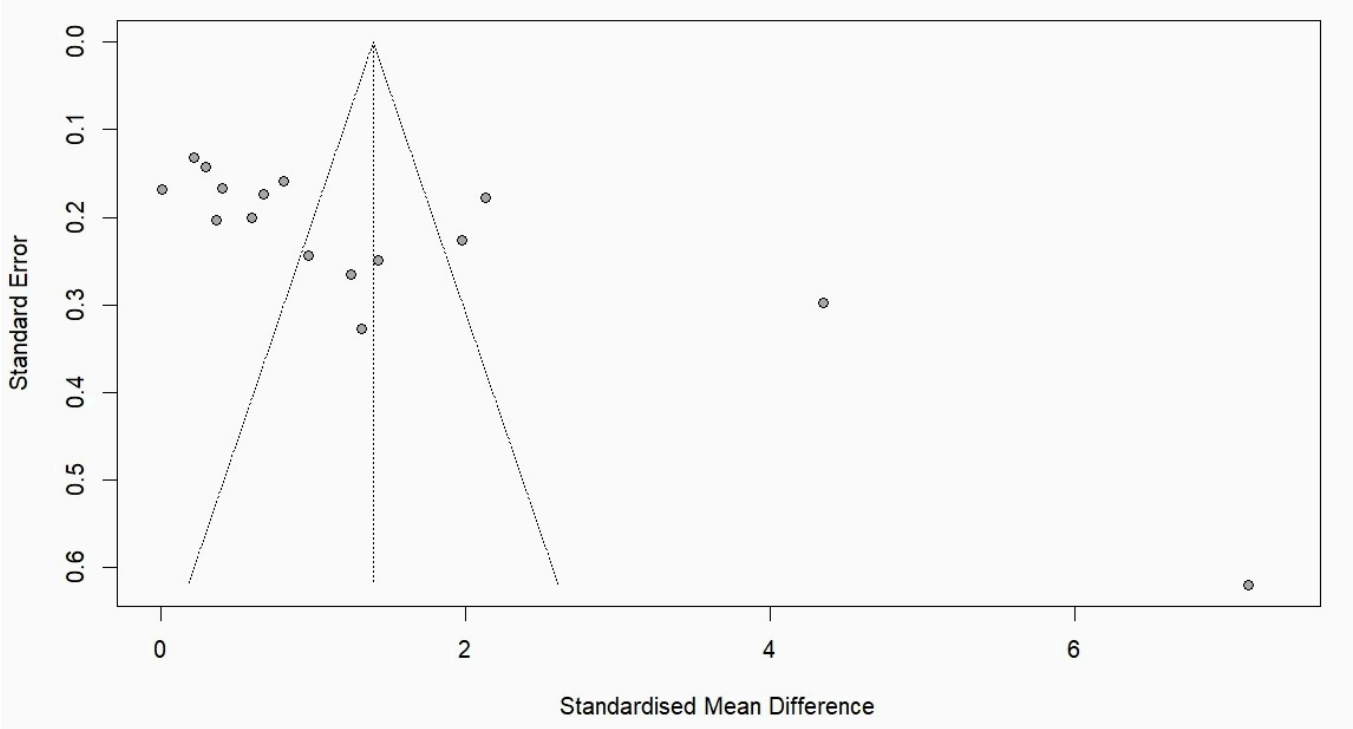
Supplementary Figure 8: Bubble plot of the meta-regression analysis evaluating the association between HbA1c and the standardized mean difference (SMD) of circulating HMGB1 levels in patients with T2DM.



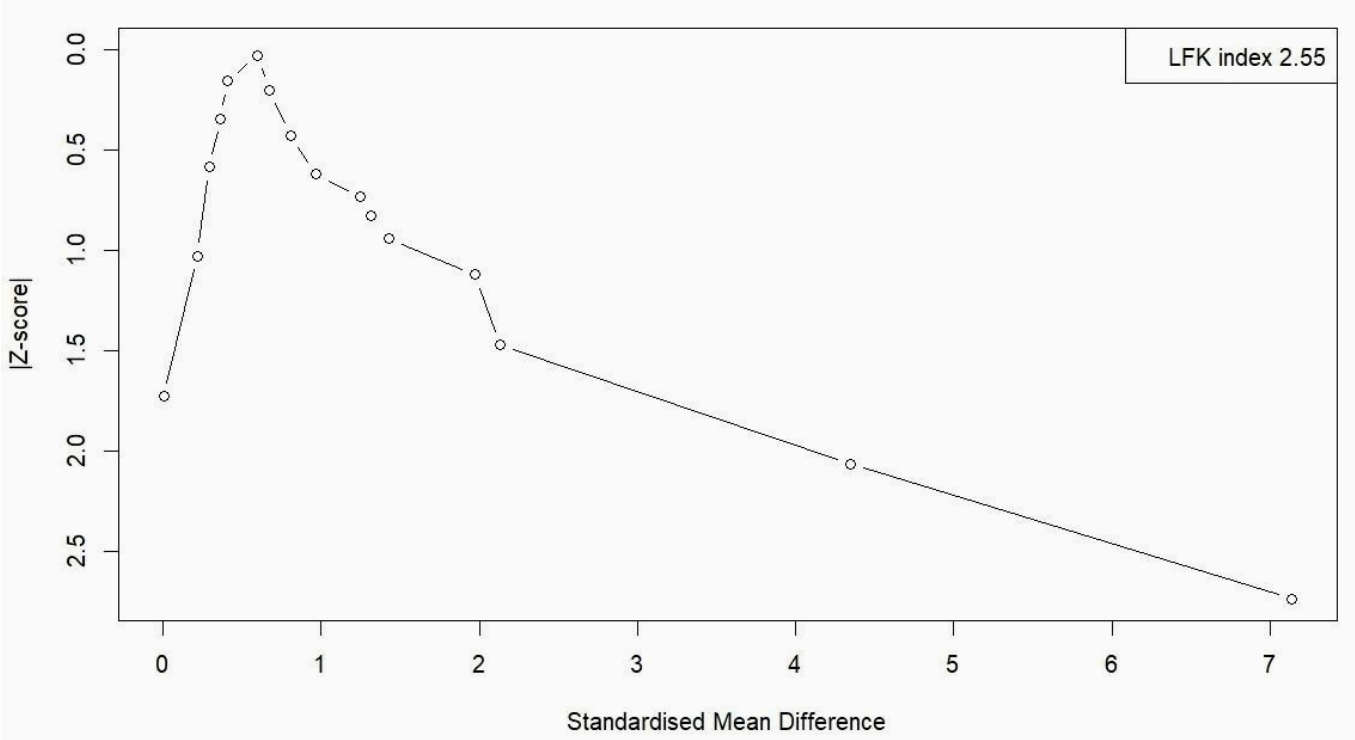
Supplementary Figure 9: Bubble plot of the meta-regression analysis evaluating the association between diabetes duration and the standardized mean difference (SMD) of circulating HMGB1 levels in patients with T2DM.



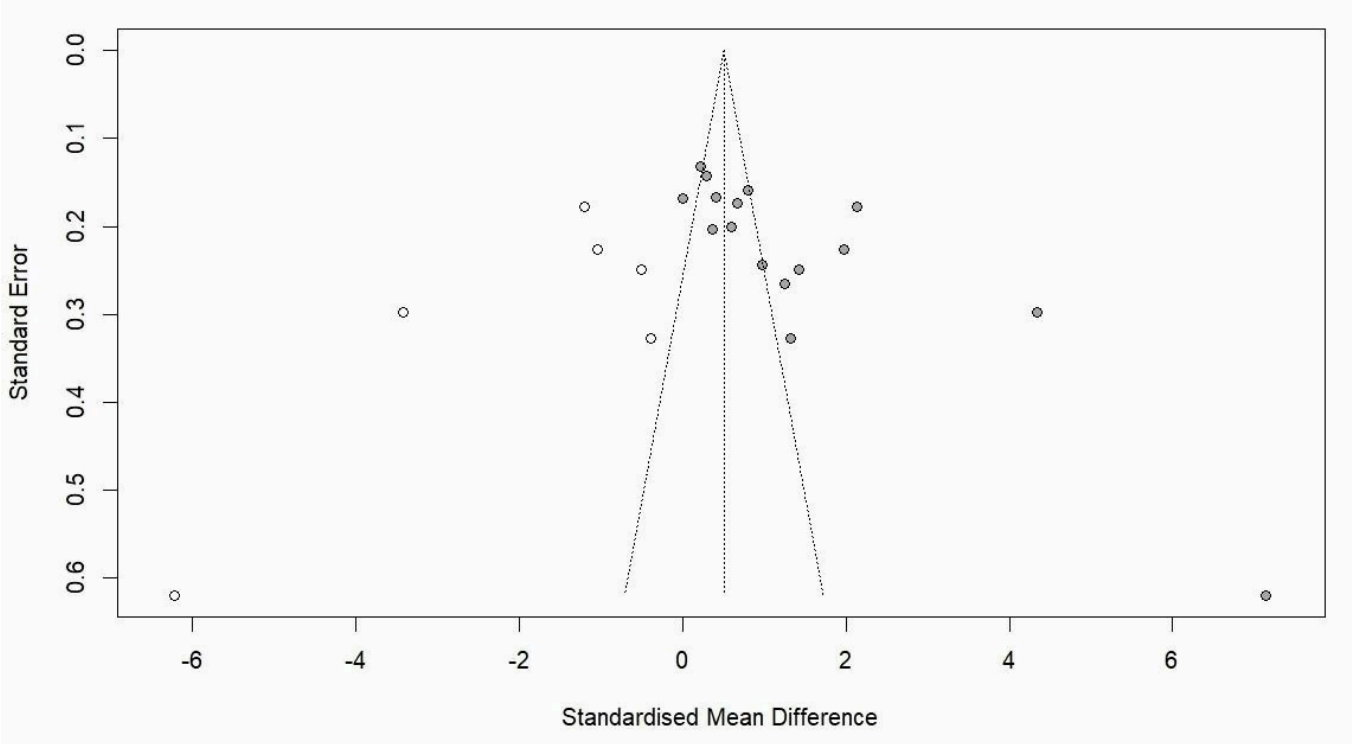
Supplementary Figure 10: Funnel plot showing publication bias in T2DM.



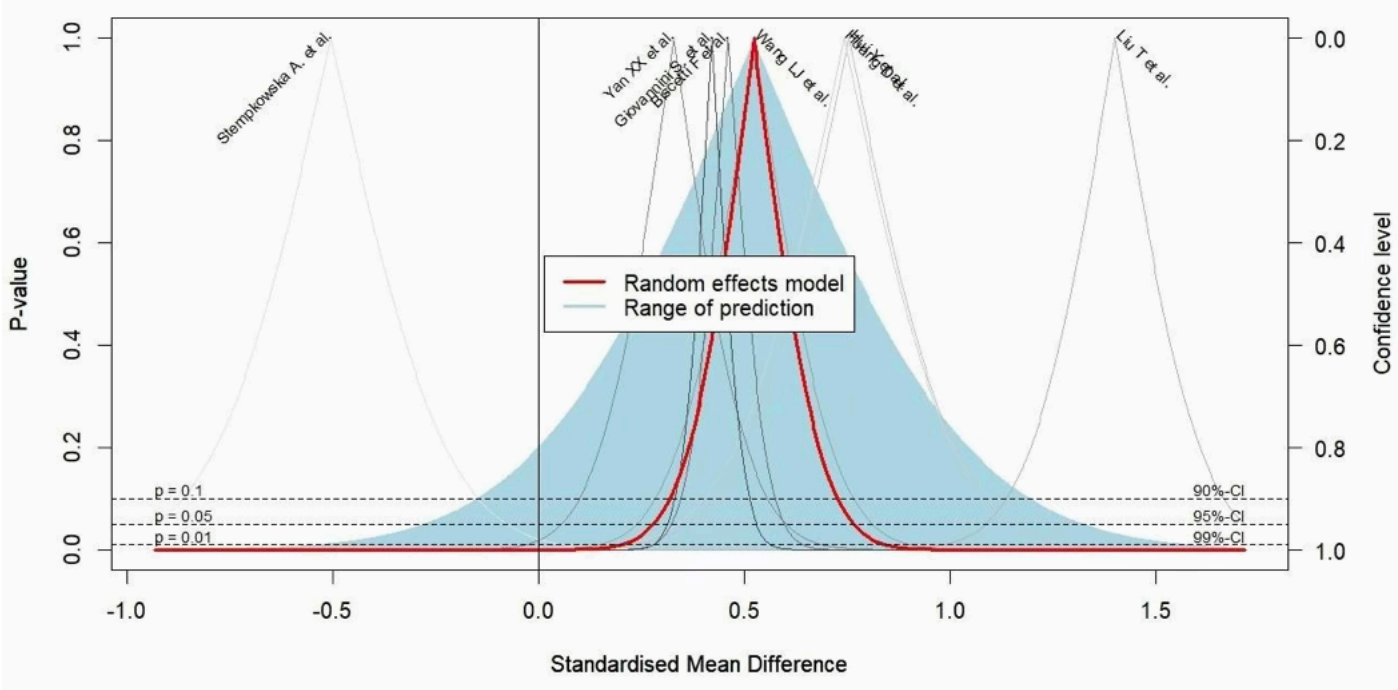
Supplementary Figure 11: Doi plot showing publication bias in T2DM.



Supplementary Figure 12: Trim-and-fill funnel plot. Six missing studies were imputed on the left side. The imputed studies are represented by circles that have no fill colour.



Supplementary Figure 13: Drapery plot depicting the meta-analysis results based on study-specific p-value functions in T2DM with complications (T2DM+C).



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

When Xanthomas Tell the Story: A Child with De Novo Familial Hypercholesterolemia

Mohnish Darshan[1], Ajay Kumar[1], Pooja Soni[1], Bhavna Dhingra[1*]

[1]*Department of Pediatrics, All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India*

Article Info

[*]***Corresponding Author:***
Bhavna Dhingra, MD, DNB
Department of Pediatrics, All India Institute of Medical Sciences, Bhopal, Madhya Pradesh, India
E-mail: bhavna.pediatrics@aiimsbhopal.edu.in

Keywords

Cutaneous xanthomas, Familial hypercholesterolemia, LDLR mutation, Pediatric dyslipidemia

Abstract

Background: Cutaneous xanthomas in children are uncommon and often overlooked, yet they may represent the earliest visible clue to severe underlying dyslipidemia. Recognising these lesions can be critical for the timely diagnosis of familial hypercholesterolemia (FH), a condition associated with lifelong cardiovascular risk.

Case Presentation: We describe a 9-year-old child presenting with multiple progressively enlarging, yellowish-orange cutaneous swellings over the gluteal region, elbows, and feet. Initially managed surgically elsewhere, these lesions' recurrent nature prompted further evaluation. Histopathology confirmed xanthomas, and lipid profiling revealed extreme hypercholesterolemia with LDL-C of 529 mg/dL. Genetic analysis identified a heterozygous pathogenic splice-site variant in the LDLR gene (c.313+1G>C), establishing the diagnosis of FH. Notably, both parents had normal lipid profiles, suggesting a likely de novo mutation. Initiation of atorvastatin therapy, combined with dietary modification, resulted in a rapid and substantial decline in LDL-C to 120 mg/dL within 1 month.

Conclusion: This case underscores that "skin can signal the heart," with xanthomas serving as early markers of severe dyslipidemia. Even in the absence of family history, FH must be considered. Early laboratory recognition and genetic confirmation enable timely intervention, significantly altering long-term cardiovascular risk.

Introduction

Familial hypercholesterolemia (FH) is an inherited lipid disorder characterised by markedly raised plasma low-density lipoprotein cholesterol (LDL-C) levels from birth. The majority of cases result from mutations in the LDL receptor gene (LDLR), while variants in apolipoprotein B (APOB) and proprotein convertase subtilisin/kexin 9 (PCSK9) are less frequent causes [1,2]. The condition is transmitted in an autosomal dominant pattern, with an estimated prevalence of heterozygous FH ranging from 1 in 200–300 individuals worldwide, and homozygous FH being much rarer (1 in a million) [2,3].

If unrecognised, lifelong exposure to elevated LDL-C accelerates atherosclerosis, conferring a 10- to 20-fold increased risk of premature coronary artery disease (CAD) [1]. While clinical events in heterozygous FH often appear in early adulthood, severe cases and homozygous FH can manifest in childhood with tendon xanthomas, xanthelasma, and even early coronary syndromes [2,4]. Case reports describe coronary artery bypass grafting and angioplasty for children under 10 years of age with homozygous FH, illustrating the aggressive natural course when left untreated [2,5].

Early identification during childhood is therefore essential. Screening based on lipid profile, combined with clinical signs such as cutaneous or tendinous xanthomas, can provide critical diagnostic clues [2].

Despite its clinical importance, FH remains substantially underdiagnosed and undertreated worldwide. In India, recent prevalence studies using Dutch Lipid Clinic Network criteria suggest FH is more common than previously estimated, but awareness remains low [3]. This gap in recognition delays intervention and leads to preventable cardiovascular morbidity. In this report, we present the case of a 12-year-old child with genetically confirmed heterozygous FH due to a de novo LDLR splice-site mutation, who presented with multiple cutaneous xanthomas and severe dyslipidemia. The absence of family history, incidental detection of additional heterozygous variants, and an excellent early response to statin therapy make this case distinct. It highlights the importance of recognising subtle cutaneous markers in childhood and stresses the need for early treatment and lifelong follow-up to avert premature atherosclerotic disease.

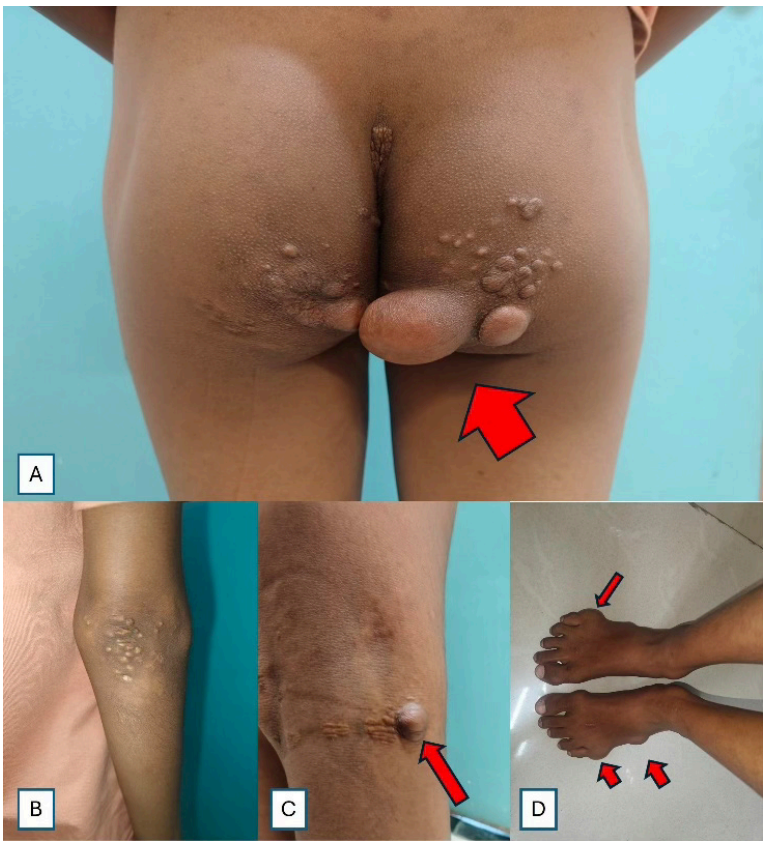
Case description

A 9-year-old child presented with multiple progressively enlarging swellings over the gluteal region for the past six months. The swellings were initially small and gradually increased in size, appearing yellowish-orange and associated with pain on walking and lying supine. The gluteal lesions were the most prominent, with the largest measuring approximately pear-sized (5 x 4 cm), soft in consistency, and smooth, without overlying skin changes or discharge. Over subsequent months, additional swellings developed over both elbows and along the dorsum and lateral borders of both feet. The elbow lesions were smaller, approximately peanut-sized (2 x 1 cm), soft, smooth, and non-tender, with no erythema, ulceration, or discharge. Some of the lesions became tender, and the child experienced significant discomfort. The patient had previously undergone multiple surgical excisions elsewhere for symptomatic relief, but the swellings recurred. No histopathological evaluation had been performed during those interventions.

The child was the fourth-born to non-consanguineous parents, with no family history of coronary heart disease, hypertension or similar skin lesions or. Birth and developmental history were unremarkable. The child was attending school with good scholastic performance. There was no history of systemic illness, drug intake, or chronic comorbidities, preceding trauma or vaccination at the site

On general examination, the child's vitals were within normal limits, and there was no pallor, icterus, cyanosis, clubbing, lymphadenopathy, or pedal oedema. Systemic examination was unremarkable. Local examination revealed multiple globular lesions over the buttocks arising from the dermis and subcutaneous tissue. The largest gluteal lesion measured approximately 5 × 4 cm, was mobile, pedunculated, soft in consistency, and had pinchable overlying skin without fluctuation or discharge. A few lesions showed mild surface abrasions with tenderness. Similar sessile lesions were noted over both elbows, with the largest measuring approximately 2 × 1 cm, soft, smooth, and non-tender, with no overlying skin changes. Examination of the feet revealed soft swellings along the lateral borders, the largest measuring 4 × 3 cm, with pinchable overlying skin, clinically resembling lipomatous lesions (Figure 1).

Figure 1: Clinical photographs showing multiple cutaneous xanthomas in a 9-year-old boy with genetically confirmed familial hypercholesterolemia.



(A) Large, globular, reddish-orange, pedunculated and sessile xanthomas over the gluteal region.

(B) Multiple sessile xanthomas over the extensor surface of the elbow.

(C) Solitary nodular lesion over the lateral aspect of the knee.

(D) Soft xanthomatous swellings along the dorsum and lateral borders of both feet (arrows).

Excision of a buttock lesion was performed, and histopathology revealed foamy macrophages within a fibrous stroma, confirming the diagnosis of xanthoma. Laboratory evaluation revealed normal complete blood counts, renal and liver function tests, and electrolyte levels. Laboratory investigations demonstrated severe dyslipidaemia with total cholesterol 647.6 mg/dL, triglycerides 76.7 mg/dL, HDL 102.8 mg/dL, and markedly elevated directly measured LDL 529.4 mg/dL. Ophthalmological evaluation did not reveal lipemia retinalis or corneal arcus.

Whole Exome Sequencing (WES) identified a heterozygous pathogenic splice-site variant in the LDLR gene (c.313+1G>C, intron 3), classified as pathogenic by ACMG criteria. This mutation disrupts the canonical donor splice site, leading to abnormal RNA splicing and loss of LDL receptor function, thereby explaining the severely elevated LDL-cholesterol levels and cutaneous xanthomas. The condition was confirmed as familial hypercholesterolemia type 1 (FHCL1; OMIM #143890). Parental lipid profiles were within normal limits, suggesting a de novo pathogenic variant; however, segregation analysis was not undertaken due to financial constraints to confirm de novo nature.

Whole exome sequencing also identified the child as a heterozygous carrier for two unrelated pathogenic variants: ABCA4 (c.5882G>A; p.Gly1961Glu), associated with Stargardt disease, and GJB2 (c.71G>A; p.Trp24*), associated with autosomal recessive nonsyndromic hearing loss. These findings are not expected to be clinically expressed in the heterozygous state due to autosomal recessive nature. Comprehensive genetic counselling was provided, addressing the implications of the LDLR variant, the possibility of mosaicism, and the incidental carrier findings.

The child was commenced on atorvastatin therapy (10mg/day) with dietary modification, and screening for cardiovascular end-organ involvement was undertaken. After one month of treatment, there was a marked improvement in the lipid profile, with serum cholesterol 189 mg/dL, triglycerides 109 mg/dL, LDL-C 120 mg/dL, and VLDL-C 21 mg/dL (Table 1).

Table 1: Lipid profile of the patient before and after initiation of atorvastatin therapy compared with normal reference ranges.

Parameters	Before atorvastatin	After atorvastatin	Normal range (mg/dL)
Cholesterol	647	189	<200
Triglycerides	76	109	<150
LDL-C	529	120	<100
VLDL-C	15.35	21	5–40
HDL	102	46	>40

Marked improvement in serum cholesterol and LDL-C (directly measured) was observed after one month of therapy.

Staged surgical excision of the residual xanthomas was planned for cosmetic correction, and the patient was advised long-term follow-up with periodic lipid monitoring, cardiovascular surveillance, ophthalmology and hearing follow-up, along with genetic counselling for the family.

Discussion

Familial hypercholesterolemia (FH) is one of the most common monogenic disorders of lipid metabolism, yet it remains under-recognised in the pediatric population. The hallmark of the disease is a persistently elevated LDL-cholesterol concentration due to pathogenic variants affecting the LDL receptor pathway [7]. The cumulative burden of high LDL from early childhood accelerates atherosclerotic changes and significantly increases the risk of premature coronary artery disease if untreated [2,8]. The presence of cutaneous or tendinous xanthomas in a child should always prompt evaluation for lipid disorders. In our patient, multiple xanthomatous swellings over the buttocks, elbows, and feet were the first clinical clues. Similar observations have been reported in other paediatric cases where tuberous or diffuse xanthomas were the presenting signs of FH [4,6]. Unfortunately, as in this case, repeated surgical excision without histopathological evaluation delayed the diagnosis. Early recognition is crucial because initiation of therapy during childhood reduces lifelong LDL exposure and improves cardiovascular prognosis [7].

Genetic testing has become central to the diagnosis of FH. Whole exome sequencing in our patient revealed a heterozygous splice-site mutation in LDLR (c.313+1G>C). Most reported paediatric FH cases involve inherited variants, either heterozygous or compound heterozygous. An unusual feature in our case was the apparent de novo occurrence of the mutation, as neither parent had a standard lipid profile. While rare, de novo mutations have been described and underline the need to consider FH even in the absence of family history [8]. Molecular confirmation not only establishes the diagnosis but also facilitates cascade screening, reproductive counselling, and individualised management strategies [7,8].
Another distinctive aspect was the incidental detection of pathogenic variants in ABCA4 (linked to Stargardt disease) and GJB2 (linked to recessive non-syndromic hearing loss). Such findings have not been commonly reported in pediatric FH cases and are clinically silent in the heterozygous state, but they broaden the genetic counselling considerations in this family.

Statins remain the first-line therapy for children with FH, with guidelines recommending initiation from 8–10 years of age if LDL-C remains ≥180 mg/dL. In our patient, atorvastatin led to a marked reduction in LDL-C within one month, consistent with the favourable responses observed in other pediatric case series. Early pharmacological treatment is supported by evidence that lipid-lowering therapy in childhood slows progression of subclinical atherosclerosis, as demonstrated by reduced carotid intima–media thickness [7,8].
Monitoring must be comprehensive and ongoing. Lipid profiles should be checked regularly, along with liver and muscle enzyme assays, to ensure the drug's safety. Non-invasive assessments such as echocardiography and carotid Doppler can detect early vascular changes. Longitudinal follow-up is essential, as highlighted by reports of children with homozygous FH developing severe coronary disease as early as the first decade of life [2].

Conclusion

This case and those previously reported share important lessons. First, childhood xanthomas are never trivial and should raise suspicion of FH. Second, genetic testing provides diagnostic certainty, even in sporadic cases without family history. Third, early statin therapy is both safe and effective in children, reducing LDL burden and improving outcomes. Finally, structured monitoring and cardiovascular surveillance are mandatory, as illustrated by cases where delayed treatment led to premature acute coronary events.

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. Goldberg AC, Hopkins PN, Toth PP, Ballantyne CM, Rader DJ, Robinson JG, Daniels SR, Gidding SS, de Ferranti SD, Ito MK, McGowan MP, Moriarty PM, Cromwell WC, Ross JL, Ziajka PE; National Lipid Association Expert Panel on Familial Hypercholesterolemia. Familial hypercholesterolemia: screening, diagnosis and management of pediatric and adult patients: clinical guidance from the National Lipid Association Expert Panel on Familial Hypercholesterolemia. J Clin Lipidol. 2011;5(3 Suppl):S1-S8. doi: 10.1016/j.jacl.2011.04.003. Epub 2011 Apr 12. PMID: 21600525.
2. Wang N, Wei Y, Zhou G, Zhang Y, Song J. Acute coronary syndrome in an 8-year-old child with familial hypercholesterolemia: a case report. Journal of Medical Case Reports. 2022;16(1):290. doi:10.1186/s13256-022-03488-3
3. Barde AK, Sethi S, Bhargav M, Waghdhare S. Indian prevalence of familial hypercholesterolemia demystified by applying Dutch lipid clinic network criteria. Int J Adv Med 2022;9:1177-1182.
4. Rahman A, Ahmed MU, Islam AKMM, Karim A, Sarker SA. A young male with familial hypercholesterolemia. J Saudi Heart Assoc. 2012;24(4):261–264. doi:10.1016/j.jsha.2012.06.264 PubMed PMID: 24174835; PubMed Central PMCID: PMC3809505.
5. Chaubal G, Pakkala AK, Nanavati AJ, Hatimi H, Sinha R, Verma L, et al. Compound Heterozygous Familial Hypercholesterolemia in a Child: Successfully Treated by Liver Transplant from Heterozygous Living Donor. Indian Pediatrics Case Reports. 2023;3(4):220. doi:10.4103/ipcares.ipcares_29_23
6. Sinha R, Singh S, Syed VS. Tuberous xanthoma in type 1 familial hypercholesterolemia. Int J Contemp Pediatr 2024;11:1317-1319.
7. Ziółkowska S, Kijek N, Zendran I, Szuster E, Barg E. Familial hypercholesterolemia – treatment update in children, systematic review. Pediatr Endocrinol Diabetes Metab. 2022;28(2):152–161. doi:10.5114/pedm.2022.116112 PubMed PMID: 35848473; PubMed Central PMCID: PMC10214937.
8. Harada-Shiba M, Ohtake A, Sugiyama D, Tada H, Dobashi K, Matsuki K, et al. Guidelines for the Diagnosis and Treatment of Pediatric Familial Hypercholesterolemia 2022. J Atheroscler Thromb. 2023;30(5):531–557. doi:10.5551/jat.CR006 PubMed PMID: 36682777; PubMed Central PMCID: PMC10164603.

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

Limitations of Transferrin- and UIBC-Based Estimation of Transferrin Saturation in Decompensated Chronic Liver Disease: A Case Series

Jagadish Ramasamy[1*]

[1]*Associate Professor and Lab Quality Manger, Department of Biochemistry, Velammal Medical College Hospital and Research Institute, Madurai, Tamil Nadu, India*

Article Info

[*]***Corresponding Author:***
Jagadish Ramasamy MD DNB
Primary Affiliation: Associate Professor and Lab Quality Manger, Department of Biochemistry, Velammal Medical College Hospital and Research Institute, Madurai, Tamil Nadu, India. Pin:625009
Current Affiliation: Assistant Professor, Department of Biochemistry, All India Institute of Medical Sciences, Madurai, Tamil Nadu, India. Pin:625009
ORCID: 0000-0003-4725-3227
E-mail: iamjagankmr@gmail.com
Phone: +91 8015597917

Keywords

Iron, unsaturated iron binding capacity, UIBC, total iron binding capacity, TIBC, transferrin saturation, TfSat, transferrin, DCLD

Abstract

Case presentation: In this case series, three patients with Decompensated Chronic Liver Disease (DCLD) were admitted to the intensive care unit for management of advanced liver dysfunction. Laboratory investigations revealed anaemia, jaundice and deranged liver function tests. A notable and consistent laboratory finding was an unquantifiably low unsaturated iron-binding capacity (UIBC), which precluded calculation of total iron binding capacity (TIBC) and transferrin saturation (TfSat) by conventional methods. To address this limitation, serum transferrin (Tf) was measured as an alternative approach to calculate TIBC. Serum Tf levels were low in all patients, yielding low calculated TIBC and elevated TfSat. However, the biochemical findings did not provide clear evidence of marked iron overload, highlighting the potential limitations of estimating TIBC from transferrin using a fixed ratio formula and, consequently, calculating TfSat in patients with markedly reduced transferrin concentrations. Therefore, calculated TfSat should be interpreted in conjunction with the clinical context and other laboratory findings when assessing iron status.

Discussion: In DCLD, reduced hepatic synthesis of Tf limits the available iron binding sites. Consequently, during UIBC estimation, minimal binding of added iron occurs, resulting in very low or unquantifiable UIBC. Under these circumstances, the use of transferrin (Tf)-based calculations as an alternative approach yielded a low TIBC and an elevated TfSat, which may not accurately reflect true iron status. Calculated TfSat may increase when the reduction in TIBC is proportionally greater than the reduction in serum iron, or when serum iron concentrations are relatively preserved despite a reduced TIBC. However, in the absence of direct TIBC measurement and an independent reference method for iron status, the relative contributions of biological and methodological factors cannot be determined.

Conclusions: This case series highlights the limitations of Tf- and UIBC-based estimation of TfSat in patients with DCLD and underscores the need for cautious interpretation of calculated iron indices in these patients.

Introduction

Decompensated Chronic Liver Disease (DCLD), also known as decompensated cirrhosis is the advanced stage of liver disease with occurrence of complications, such as ascites, hepatic encephalopathy, gastrointestinal bleeding and portal hypertension [1]. DCLD and iron homeostasis are closely interlinked making interpretation of routine iron status markers particularly challenging and potentially misleading [2].

Iron is essential for the synthesis of heme (a component of hemoproteins) and as a cofactor for enzymes [3]. Markers of body iron status include serum iron, ferritin (Ftn), total iron-binding capacity (TIBC), serum transferrin and transferrin saturation (TfSat) [4]. Serum ferritin is indicative of body iron stores; its concentrations are decreased in the presence of iron deficiency and increased in conditions of iron overload. However, it is also an acute phase reactant, with concentrations being increased in the presence of inflammation [5]. Transferrin (Tf) is the major iron transporter in blood; its capacity to bind iron is expressed as total iron-binding capacity (TIBC). TIBC can be estimated using direct or indirect methods. In labs in India, indirect methods are widely used and formulae were employed to calculate TIBC either by from unsaturated iron-binding capacity (UIBC) or Tf values. The UIBC-based approach is commonly used in diagnostic laboratories in country like India due to its cheaper cost and wide availability. Despite its widespread use, it is subjected to errors. The Tf-based calculation of TIBC is rarely employed due to high cost, inadvertent errors in unit conversion and use of appropriate correction factor [4,5]. Kits are available for direct TIBC estimation, but their widespread use is limited due to a shortage of manufacturers and standardisation.

In this case series, three patients presented with a distinct laboratory finding - their serum unsaturated iron-binding capacity (UIBC) was too low to be quantified; as a result, it was not possible to determine values for TIBC and TfSat. All three patients had chronic liver disease and were advised to undergo liver transplantation.

Case Presentation

The detailed case presentations are as follows.

Case 1

A 51-year-old man, known to have alcohol-related DCLD with portal hypertension, was admitted in the intensive care unit (ICU) due to complaints of fever and drowsiness. He was on the cadaveric liver transplantation list, as part of recommended treatment. He is a non-smoker and does not have any other co-morbidities. Upon examination, he was found to be icteric and pyrexia was observed. Examination of vital signs showed presence of tachycardia (heart rate 138/minute) and hypotension (100/50 mm Hg). The patient was administered intravenous fluids, antibiotics, antifungals and antiemetics. Results of his laboratory investigations are shown in Table 1. He underwent conservative management in the ICU for five days. His general condition improved; he was discharged on sixth day from ICU to ward with the recommendation for follow up in the transplant clinic.

Case 2

A 43-year-old man was admitted in the emergency room with complaints of fever, loss of weight for the past one week. Upon examination, he was found to be icteric and pyrexia was observed. He was diagnosed to have type-2 diabetes mellitus a month ago. He is a smoker and has been an alcoholic for a decade and does not have any other co-morbidities. Examination of vital signs revealed a heart rate of 90/minute and a blood pressure reading of 110/60 mm Hg. Ultrasound and computed tomography scans showed evidence of chronic liver disease with portal hypertension and right pyelonephritis. Results of his laboratory investigations are shown in Table 1. He was managed conservatively with antacids, antiemetics, and antibiotics. He was diagnosed to have DCLD with portal hypertension and advised to undergo liver transplantation as the further course of management. His general condition improved, and he was discharged with the recommendation for follow up in the transplant clinic.

Case 3

A 48-year-old man, known to have DCLD, was admitted in the intensive care unit (ICU) with presenting complaints of abdominal distension, decreased urine output, itching and constipation. He had raised similar complaints two years ago and was advised regular follow-up for further management. However, he was irregular with follow-up and also attempted native medicines for medical management for three months. He is a alcoholic for two decades and does not have any other co-morbidities. Upon examination, he was found to have bilateral pitting pedal edema and a moderate degree of ascites with distended veins in the abdomen. Examination of vital signs revealed a heart rate of 80/minute and blood pressure of 110/70 mm Hg. He was administered intravenous fluids, antibiotics, antifungals and antiemetics. Results of his laboratory investigations are shown in Table 1. He was advised to undergo liver transplantation as the further course of management.

Table 1: Laboratory investigations.

	Reference range	Case 1	Case 2	Case 3
Hemoglobin, g/dL	13-17	9.1 ↓	9.2 ↓	9.4 ↓
Mean corpuscular volume (MCV), fL	83-101	99.9	98.3	99.5
Mean corpuscular hemoglobin (MCH), pg	27-32	37.7 ↑	35.2 ↑	36.2 ↑
Mean corpuscular hemoglobin concentration (MCHC), g/dL	31-34	37.7 ↑	35.8 ↑	36.4 ↑
Red cell distribution width (RDW), %	11-14	13.8	13.1	14.3 ↑
Total bilirubin, mg/dL	0.3-1.2	37.41 ↑	2.22 ↑	7.92 ↑
Indirect Bilirubin, mg/dL	0.2-1	18.87 ↑	0.72	7.13 ↑
Direct Bilirubin, mg/dL	0-0.3	18.54 ↑	1.50 ↑	6.19 ↑
Total protein, g/dL	6-7.8	5.1 ↓	7.1	6.3
Albumin, g/dL	3.5-5.5	2.5 ↓	3.1 ↓	2.7 ↓
Globulins, g/dL	2.5-3.5	2.6	4	3.6
A:G ratio	1.2-1.5:1	0.96 ↓	0.77 ↓	0.75 ↓
AST, IU/L	10-40	263 ↑	43 ↑	42 ↑
ALT, IU/L	10-40	189 ↑	26	19
ALP, IU/L	36-92	118 ↑	57	133 ↑
Prothrombin time (PT), in seconds	10-13	38 ↑	18 ↑	22 ↑
PT-INR	0.8-1.1	3.2 ↑	1.5 ↑	1.9 ↑
MELD 3.0 Score	6–40 (higher scores indicate worse prognosis)	38	18	27
Child Pugh Score	5-15 Class A (5–6), Class B (7–9), and Class C (10–15)	11	7	12
Serum urea, mg/dL	17-43	104 ↑	31	19
Serum creatinine, mg/dL	0.7-1.4	1.4	0.7	0.9
Hepatitis B Surface Antigen (HBsAg)	Signal-to-cutoff (S/CO) ratio: <1.0 - Negative, ≥1.0 - Positive	Negative	Negative	Negative
Anti-Hepatitis B Surface Antibody titre, mIU/mL (Anti-HBs)	<10 - Non-protective >10 - Protective	Not done	Not done	0.1
Anti-Hepatitis C Virus Antibody (Anti-HCV)	Signal-to-cutoff (S/CO) ratio: <1.0 - Negative, ≥1.0 - Positive	Negative	Negative	Negative
Hepatitis B Core Antibody (Anti-HB Core)	Signal-to-cutoff (S/CO) ratio: <1.0 - Negative, ≥1.0 - Positive	Not done	Not done	Negative
HIV-1 and 2 (Antigen and Antibody)	Signal-to-cutoff (S/CO) ratio: <1.0 - Non-reactive, ≥1.0 - Positive	Non-reactive	Non-reactive	Non-reactive
Serum iron, µg/dL	59-158 (males) 37-145 (females)	55 ↓	174 ↑	150
Unsaturated iron-binding capacity (UIBC), µg/dL	112-347	<17	<17	<17
Serum total iron-binding capacity (TIBC), µg/dL (calculated from UIBC)	250-400	Indeterminable	Indeterminable	Indeterminable
Serum ferritin, ng/mL	20-250 (males) 10-120 (females)	475 ↑	924 ↑	159
Transferrin saturation (TfSat), % (calculated using iron and TIBC)	14-50	Indeterminable	Indeterminable	Indeterminable

↑ and ↓ denote values above and below the respective reference ranges. Complete blood count (CBC) was performed using the Beckman Coulter DxH 800 analyzer. PT-INR was measured using the ACL Elite Pro analyzer. Serum iron and UIBC were measured by the ferrozine colorimetric method, in which UIBC was determined from the difference between added and unbound ferrous iron, using a Mindray UIBC reagent kit on a Mindray BS-430 autoanalyzer. Serum liver function tests, urea, and creatinine were measured using colorimetric assays with Agappe reagents on a Toshiba 120-FR autoanalyzer. Serum ferritin was measured by latex-enhanced immunoturbidimetry using Agappe ferritin reagent on a Toshiba 120-FR autoanalyzer. Viral serology (HBsAg, anti-HBs, anti-HBc, anti-HCV, and HIV-1/2 antigen/antibody) was performed using a chemiluminescence immunoassay on the Vitros 3600 analyzer.

Case work-up

Firstly, pre-analytical errors such as timing of sample collection, HIL (hemolysis, icterus, lipemia) indices of sample was documented and recent intake of iron supplementation and intravenous iron administration were ruled out. Each sample was rerun after serial pre-dilution and re-analysed to rule out any matrix effect, results of UIBC were unchanged. In addition, according to the manufacturer's specifications, the UIBC assay is not affected by bilirubin concentrations up to 40 mg/dL. The highest bilirubin concentration observed in this case series was 37.4 mg/dL, indicating that bilirubin interference was unlikely.

In routine lab practice, TIBC and TfSat are calculated using serum iron and UIBC, using mathematical formulae shown below [4]. Since the value of UIBC was not quantifiable, it was not possible to use these formulae to calculate values for TIBC and TfSat.

$$TIBC, \mu g/dL = Serum\ iron\ (in\ \mu g/dL) + Serum\ UIBC\ (in\ \mu g/dL)$$

$$TfSat, \% = \left(\frac{Serum\ iron\ (in\ \mu g/dL)}{Serum\ TIBC\ (in\ \mu g/dL)}\right) * 100$$

An alternative method to calculate TIBC and TfSat is by using serum transferrin [4]. Therefore, serum transferrin was directly measured, and TIBC and TfSat were subsequently calculated in these three patients, using the following formulae:

$$TIBC, in\ \mu g/dL = 1.41 * Serum\ transferrin\ (in\ mg/dL)$$

$$TfSat, \% = \left(\frac{Serum\ iron\ (in\ \mu g/dL)}{Serum\ transferrin\ (in\ mg/dL) * 1.41}\right) * 100$$

The calculated values for TIBC and TfSat are shown in Table 2. All three patients had low transferrin and total iron-binding capacity; however, their TfSat values were very high (Table 2).

Table 2: Transferrin-based TIBC and TfSat.

	Reference range	Case 1	Case 2	Case 3
Serum transferrin, mg/dL	200-360	42 ↓	98 ↓	131 ↓
Serum TIBC, µg/dL (calculated from transferrin)	250-400	59 ↓	137 ↓	184 ↓
Serum TfSat, % (calculated using iron and transferrin)	14-50	93 ↑	126 ↑	81 ↑

↑ and ↓ denote values above and below the respective reference ranges. Serum Transferrin was measured using immunoturbidimetric method using Roche e411autoanalyser.

Case discussion

UIBC-based approach

The method used to estimate UIBC explains the low values seen in patients with chronic liver disease. UIBC is measured by saturating transferrin in serum at an alkaline pH through addition of known excess amount of iron. The unbound ferrous ions were subsequently measured using the ferrozine method. The difference between the amount of iron added and the unbound iron measured is taken as the UIBC of the sample.

UIBC,(g/dL)= Amount of ferrous iron added -Amount of unbound ferrous iron

TIBC is then calculated as described earlier by adding the UIBC to the serum iron concentration. This approach is widely used in labs, and does not require a separation step to remove non-transferrin bound iron, unlike direct TIBC estimation methods [4]. However this method is prone to analytical variability arising from iron contamination, as well as differences in calibrators and control materials [6]. A previous study have demonstrated a negative bias in the estimation of both TIBC and TfSat using this method [7].

In those with advance chronic liver disease, transferrin concentrations are reduced due to impaired synthetic function of liver. The reduced transferrin results in diminished capacity to bind the added iron. Consequently, the calculated difference between the added and unbound iron—that is, the UIBC—is markedly low. Given the analytical reporting range of the assay (17–558 µg/dL), UIBC values in these patients were reported as <17 µg/dL. As the UIBC was below the quantifiable range and an absolute value could not be obtained, TIBC and transferrin saturation could not be reliably calculated in these cases. By substituting the UIBC value as "17 µg/dL", TfSat would be approximately 76%, 91% and 90% suggesting the possibility of iron overload.

Tf-based approach

This is an alternative, less widely employed approach to calculate TIBC and TfSat and formula has been described earlier [4]. The serum TfSat calculated by Tf-based approach were higher in all three patients, suggesting the possibility of iron overload, if interpreted in isolation (Table 2).

Transferrin (Tf), predominantly synthesised in the liver is a major iron transporter. It carries two iron atoms in ferric state (di-ferric Tf) and its levels are determined by the systemic iron status and largely by the biosynthetic function of the liver. Hence, in patients with DCLD due to impaired synthetic function, transferrin and UIBC levels were decreased. As TfSat is calculated using serum iron as the numerator and TIBC as the denominator, a disproportionate reduction in TIBC may result in elevated calculated TfSat despite the absence of convincing biochemical evidence of marked iron overload. This may occur when the reduction in TIBC is proportionally greater than the reduction in serum iron or when serum iron concentrations remain relatively preserved despite reduced TIBC.

Among the three patients, serum iron levels were below the reference range in patient 1, mildly increased in patient 2, and within the reference range in patient 3. Serum ferritin levels were elevated in patients 1 and 2, while remaining within the reference range in patient 3. In DCLD, as hepatic transferrin synthesis declines, iron exceeds the reduced binding capacity of transferrin and shifts into the non-transferrin bound iron (NTBI) pool. The standard ferrozine colorimetric assay for serum iron measures both iron bound to transferrin (TBI) and NTBI [8], hence serum iron concentrations may remain relatively preserved or elevated despite reduced transferrin levels.

Alcohol-related decompensated liver disease was the underlying etiology in all three patients, and alcohol consumption has been shown to perturb iron homeostasis [9]. The Child-Pugh score classified Patient 1 and 3 as Class C (decompensated cirrhosis) and Patient 2 as Class B (significant functional compromise). The high MELD 3.0 scores further indicate advanced liver disease and a poorer prognosis. These findings reemphasise the marked impairment of hepatic synthetic function, resulting in reduced transferrin production and consequently lower UIBC/TIBC levels. A study have shown low transferrin correlated with poor outcome and predict survival in patients with liver disease awaiting transplantation [10]. All three patients in this case series were awaiting liver transplantation. Therefore, in patients with advanced DCLD, calculated TIBC and TfSat should be interpreted together with serum iron, ferritin, and the overall clinical context, rather than in isolation, when assessing iron status.

There are other markers that can be employed to assess iron status in those with DCLD. The studies have reported the diagnostic utility of soluble transferrin receptor (sTfR):log ferritin ratio in differentiating iron deficiency anemia (IDA) and anemia of chronic disease (ACD) [11].

Limitations

A major limitation of this study was direct TIBC estimation was not done, hence a method specific bias cannot be excluded. Hence, it is not possible to determine whether the observed elevation in TfSat is attributable to the underlying disease process, limitations of the UIBC methodology, or the mathematical conversion factors used to derive TIBC and TfSat. Furthermore, the absence of an independent reference method, such as MRI-based liver iron quantification or liver biopsy, limits our ability to definitively establish the true iron status of these patients. Inflammatory markers such as C-reactive protein and procalcitonin were not measured; therefore, the potential contribution of inflammation to the observed ferritin levels could not be assessed.

Conclusion

This case series highlights the limitations of transferrin- and UIBC-based estimation of transferrin saturation (TfSat) in patients with decompensated chronic liver disease (DCLD). These patients may present with low transferrin concentrations, low UIBC/TIBC values, and elevated calculated TfSat. In this setting, elevated TfSat should be interpreted cautiously and in conjunction with supportive clinical and laboratory evidence, as it may not accurately reflect true iron status. Overall, an integrated approach to iron evaluation is essential in DCLD to avoid diagnostic errors and support appropriate clinical decision-making.

Statement and Declarations

Funding sources

This study does not require any funding.

Declaration of interests

None.

Ethical Approval

This case series utilized anonymized patient data and was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval was waived by the Institutional Ethics Committee.

Credit Author Statements

JR: Study Design, Conceptualization, Writing – Original Draft, Reviewing And Editing, Data Visualization, Validation, Methodology, Formal Analysis And Supervision.

References

1. D'Amico G, Bernardi M, Angeli P. Towards a new definition of decompensated cirrhosis. J Hepatol 2022;76:202–7 DOI: https://doi.org/10.1016/j.

jhep.2021.06.018
2. Gkamprela E, Deutsch M, Pectasides D. Iron deficiency anemia in chronic liver disease: etiopathogenesis, diagnosis and treatment. Ann Gastroenterol 2017;30:405–13 DOI: https://doi.org/10.20524/aog.2017.0152
3. Kennelly PJ, Murray RK, Jacob M, Varghese J. Plasma Proteins & Immunoglobulins. In: Rodwell VW, Bender DA, Botham KM, Kennelly PJ, Weil PA, editors. Harpers Illus. Biochem. 31st ed., New York, NY: McGraw-Hill Education; 2018
4. Burtis CA, Ashwood ER, Bruns DE. Tietz Textbook of Clinical Chemistry and Molecular Diagnostics - E-Book: Chapter 32 - Hemoglobin, Iron and Bilirubin. Elsevier Health Sciences; 2011
5. Pfeiffer CM, Looker AC. Laboratory methodologies for indicators of iron status: strengths, limitations, and analytical challenges. Am J Clin Nutr 2017;106:1606S-1614S DOI: https://doi.org/10.3945/ajcn.117.155887
6. Lo S. Discontinuation of calculated TIBC and UIBC, and change in grading status for percent Iron Saturation 2007
7. Blanck HM, Pfeiffer CM, Caudill SP, Reyes M, Gunter EW, Imperatore G, et al. Serum iron and iron-binding capacity: a round-robin interlaboratory comparison study. Clin Chem 2003;49:1672–5 DOI: https://doi.org/10.1373/49.10.1672
8. Ito S, Ikuta K, Kato D, Shibusa K, Niizeki N, Tanaka H, et al. Non-transferrin-bound iron assay system utilizing a conventional automated analyzer. Clin Chim Acta 2014;437:129–35 DOI: https://doi.org/10.1016/j.cca.2014.07.013
9. Ali N, Ferrao K, Mehta KJ. Liver Iron Loading in Alcohol-Associated Liver Disease. Am J Pathol 2023;193:1427–39 DOI: https://doi.org/10.1016/j.ajpath.2022.08.010
10. Viveiros A, Finkenstedt A, Schaefer B, Mandorfer M, Scheiner B, Lehner K, et al. Transferrin as a predictor of survival in cirrhosis. Liver Transpl 2018;24:343–51 DOI: https://doi.org/10.1002/lt.24981
11. Marković M, MajkićSingh N, Subota V. Usefulness of soluble transferrin receptor and ferritin in iron deficiency and chronic disease. Scand J Clin Lab Invest 2005;65:571–6 DOI: https://doi.org/10.1080/00365510500206542

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

Hyperferritinemia and Beyond: Biochemical Insights from a Case Series of Secondary Hemophagocytic lymphohistiocytosis (HLH)

Lekha Priyadharshini Kamarajan[1], Sushil Kumar[1], Ravikirti[1] , Janmeshwar Prasad[1] , Mala Mahto[1*]

[1]*AIIMS Patna, Phulwari Sharif, Patna, Bihar, India*

Article Info

[*]***Corresponding Author:***
Mala Mahto, MD Biochemistry, Professor, Biochemistry, AIIMS Patna, Phulwari Sharif, Patna, Bihar 801507, India
ORCID: https://orcid.org/ 0000-0003-0445-6972
E-mail: dr.malamahto@gmail.com

Keywords

Hemophagocytic Lymphohistiocytosis, Systemic Lupus Erythematosus, Autoimmune Hemolytic Anemia, Severe Dengue, Autoimmune Disease, fever, hyperferritinemia

Abstract

Hemophagocytic Lymphohistiocytosis (HLH) is a rare, life-threatening hyperinflammatory syndrome characterized by excessive immune activation, classified as either primary (familial) or secondary, triggered by infections, malignancies, autoimmune diseases or immune deficiencies. Manifestations include persistent fever, hepatosplenomegaly, cytopenias and elevated ferritin levels. Biochemical markers play a critical role in diagnosing HLH with elevated ferritin, hypofibrinogenemia and hypertriglyceridemia. Three adolescent female cases of secondary HLH highlight the laboratory's critical role in diagnosis and management. In the first (Pseudomonas infection), elevated ferritin, liver enzymes, and low fibrinogen guided antibiotic, steroid, and etoposide therapy. The second, with systemic lupus erythematosus (SLE) and autoimmune hemolytic anemia (AIHA), was diagnosed using elevated ferritin, liver enzymes, and autoimmune markers, managed with hydroxychloroquine and steroids. The third (severe dengue) showed hyperferritinemia, elevated liver enzymes, and hypofibrinogenemia, guiding steroid, iron, and transfusion therapy. These cases highlight the role of biochemical markers in early diagnosis of HLH.

Introduction

Secondary Hemophagocytic Lymphohistiocytosis (HLH) is a critical hyperinflammatory condition that occurs in response to infections, autoimmune disorders, or malignancies [1]. In contrast to the genetic basis of primary HLH, secondary HLH arises due to excessive immune activation from external factors, leading to a cytokine storm and subsequent multi-organ involvement. Patients typically present with persistent fever, hepatosplenomegaly, and cytopenias, often mimicking sepsis or other inflammatory diseases, which makes early recognition difficult. Laboratory parameters, including raised ferritin and liver enzymes, hypertriglyceridemia, and low fibrinogen levels, play a pivotal role in diagnosis [2,3]. This case series discusses three adolescent female patients who developed secondary HLH due to different underlying causes, including systemic autoimmune disease, bacterial infection, and severe dengue infection. These cases emphasize the need to correlate clinical features with laboratory findings for prompt diagnosis and management of HLH, which is vital for improving patient outcomes.

Case series

Case 1

A 16-year-old female presented with a three-month history of fever (102°F), generalized weakness for ten days, itching for two days, and vomiting for two days. She reported migratory joint pains affecting large joints but denied evening fevers, weight loss, or urinary symptoms. On examination, she was pale with multiple enlarged lymph nodes, was fully oriented, and had no neurological deficits. Her abdomen was soft and non-tender with clear breath sounds. Laboratory evaluation revealed anemia, leukocytosis, raised inflammatory markers, and a mild pleural effusion (Table 1). Blood cultures grew Pseudomonas aeruginosa, prompting initiation of Piperacillin-Tazobactam, later escalated to Meropenem due to persistent fever and subsequent leukopenia. Bone marrow biopsy (Figure 1a and Figure 1b) demonstrated hemophagocytosis, while lymph node biopsy showed reactive hyperplasia. Imaging confirmed hepatosplenomegaly and pleural effusion. Elevated ferritin and LDH levels supported the diagnosis of secondary Hemophagocytic Lymphohistiocytosis (HLH). Tropical infections including malaria and dengue were excluded based on clinical assessment, negative serology, and the absence of typical features such as thrombocytopenia or marked liver dysfunction. The diagnosis of secondary HLH was established according to the HLH-2004 criteria based on persistent fever, splenomegaly, cytopenias involving two cell lineages, hypofibrinogenemia, hyperferritinemia, and bone marrow evidence of hemophagocytosis. Management included Meropenem for infection control and dexamethasone for HLH treatment. The patient was advised to monitor her temperature every 4 hours and was prescribed paracetamol 650 mg (maximum four tablets per day), pantoprazole 40 mg once daily before breakfast, calcium once daily, and dexamethasone 6 mg in the morning and 4 mg in the evening, with tapering planned on an outpatient basis.

Figure 1a and Figure 1b: demonstrated hemophagocytosis, while lymph node biopsy showed reactive hyperplasia.

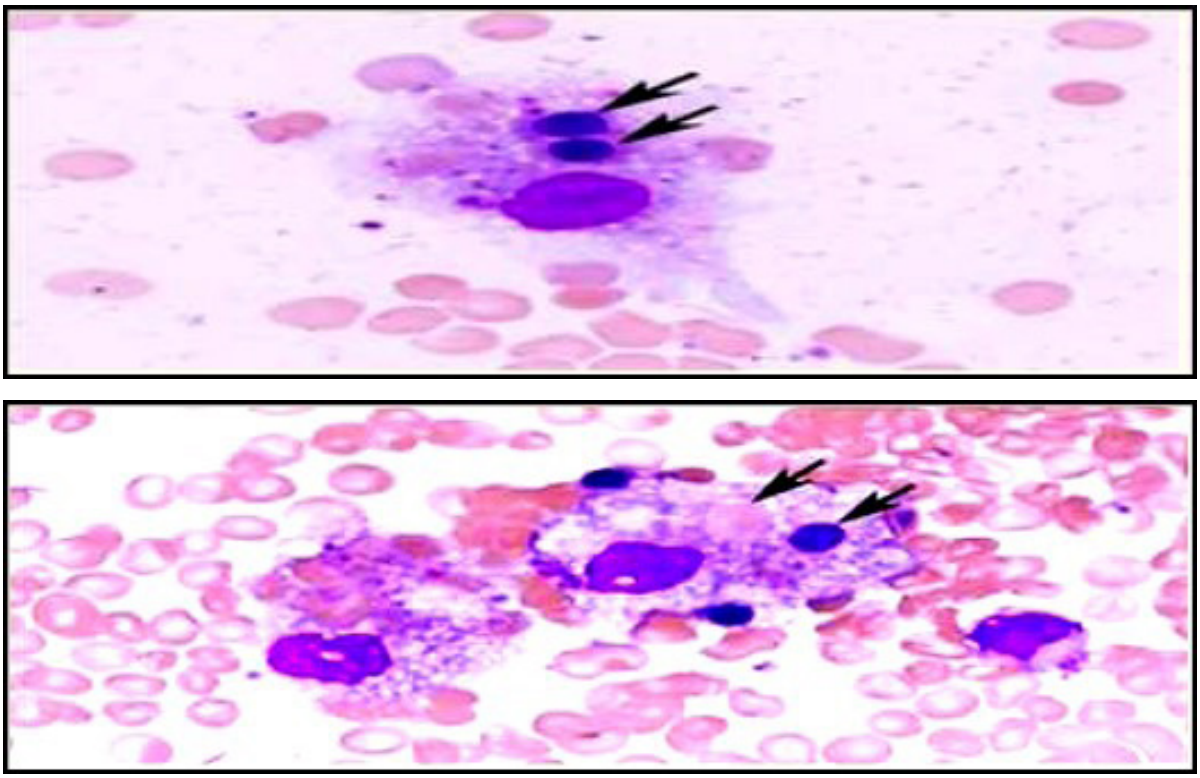


Table 1: Laboratory parameters, imaging, and autoimmune profiles assessed in patients with secondary hlh, aiding diagnosis and monitoring. Values reflect baseline findings prior to initiation of hlh-specific therapy.

Lab Parameter	Normal Range	Case 1	Case 2	Case 3	Machine/Company (Origin)
Hemoglobin (g/L)	120-160	70	76	99	Sysmex XN 1000 (Sysmex Corporation, Japan)
Leukocyte Count (/µL)	4,000-11,000	22,000	9,800	9,900	Sysmex XN 1000 (Sysmex Corporation, Japan)
Platelet Count (× 10^3/µL)	150-450	50	60	120	Sysmex XN 1000 (Sysmex Corporation, Japan)
C-Reactive Protein (mg/L)	<10	212	132	72	Beckman Coulter AU680 (Beckman Coulter, USA)
Ferritin (µg/L)	10-291	>1000	>1500	26,868	Siemens ADVIA Centaur XP (Siemens Healthineers, Germany) & Vitros 5600 (Ortho Clinical Diagnostics, USA)
Fibrinogen (mg/dL)	200-400	50	80	49	Siemens ADVIA Centaur XP (Siemens Healthineers, Germany)
Lactate Dehydrogenase (LDH, U/L)	100-250	1300	600	1512	Beckman Coulter AU680 (Beckman Coulter, USA)
Urinary Microalbumin Creatinine Ratio (µg/mg)	<30	657.44	300	500	Beckman Coulter AU680 (Beckman Coulter, USA)
Direct Coombs Test (DCT)	Negative	Positive	Positive (4+)	Positive (4+)	Bio-Rad
Blood Culture	Negative	Pseudomonas aeruginosa	Negative	Negative	BD BACTEC FX (Becton Dickinson, USA)
Differential Leukocyte Count (%)	A 40-80, L: 20-40, M 2-10, E: 1-6, B: 0-1	N 66.9, L 29.7, M 2.9, E 0.4, B 0.1	N 66.9, L 28.2, M 3.8, E 0.9, B 0.2	N 77.5, L 18.1, M 3.1, E 1.0, B 0.3	Sysmex XN 1000 (Sysmex Corporation, Japan)
Absolute WBC Differentials count (/µL)	–	N: 14,718, L: 6,534, M: 638, E: 88, B: 22	N: 6,556, L: 2,764, M: 372, E: 88, B: 20	N: 7,673, L: 1,792, M: 307, E: 99, B: 30	Calculated by formula
Total Bilirubin (µmol/L)	<20.52	11.2	37.6	117.9	Beckman Coulter AU680 (Beckman Coulter, USA)
Direct Bilirubin (µmol/L)	<5.13	3.42	20.52	99.18	Beckman Coulter AU680 (Beckman Coulter, USA)
ALT (U/L)	10-40	–	111	562	Beckman Coulter AU680 (Beckman Coulter, USA)
AST (U/L)	10-40	–	273	1512	Beckman Coulter AU680 (Beckman Coulter, USA)
Triglycerides (mmol/L)	<1.69	–	5.6	3.3	Beckman Coulter AU680 (Beckman Coulter, USA)
Procalcitonin (µg/L)	<0.5	–	3.0	2.2	Siemens ADVIA Centaur XP (Siemens Healthineers, Germany)
PT (seconds)/INR	9.9-13.0 / 0.8-1.2	–	–	31.6/2.53	Stago STA Compact Max (Stago, France)
Ultrasound abdomen & Chest X-ray Findings	–	Mild pleural effusion	Bilateral mild pleural effusion & splenomegaly	Gross ascites & bilateral mild pleural effusion	Siemens Ysio Max (Siemens Healthineers, Germany)
CECT Scan Findings	–	Hepatosplenomegaly, bilateral pleural effusion	–	–	GE Revolution CT (GE Healthcare, USA)
ANA Pattern	Negative	–	Homogeneous (4+)	–	IIFT (EUROIMMUN, Germany)

PT: Prothrombin time, INR: International normalised ratio, CECT; Contrast enhanced Computed tomography, ANA: anti-nuclear antibodies, IIFT: Indirect immunofluorescence test, ALT: alanine aminotransferase, AST: Aspartate transaminase, LDH: Lactate Dehydrogenase.
Ferritin concentrations in Cases 1 and 2 exceeded the analytical measuring range of the assay and were therefore reported as >1000 ng/mL and >1500 ng/mL, respectively. These values represent the lower limit above the assay's measurable range and should not be interpreted as the actual serum ferritin concentrations.

Case 2

An 18-year-old female presented with a two-month history of intermittent fever, oral ulcers, and joint pain. The fever, accompanied by chills, responded to intravenous paracetamol. Oral ulcers, initially located on the hard palate, progressed to the ventral tongue, causing difficulty in swallowing. Joint pains started in the proximal interphalangeal joints and later affected the ankles and knees. Her history included abnormal body movements eight years prior and oral ulcers three years ago. On examination, she was conscious and oriented but had hypotension (96/52 mmHg) and a high-grade fever (105°F). Respiratory examination revealed bilateral wheezing, and splenomegaly was noted on abdominal examination. Initial laboratory investigations showed anemia, leukopenia, thrombocytopenia, raised C-reactive protein (CRP), and abnormal liver function tests (Table 1). Autoimmune evaluation confirmed systemic lupus erythematosus (SLE) with a positive ANA and positivity for SmD1: 3+, dsDNA, Histone, SSA/Ro-60, CENP-B: 1+, and equivocal results for AMA M2 and Nucleosome antibodies. A strongly positive Direct Coombs

Test (4+) indicated autoimmune hemolytic anemia (AIHA). A diagnosis of SLE complicated by secondary Hemophagocytic Lymphohistiocytosis (HLH) and AIHA was made. The diagnosis of secondary HLH was established according to the HLH-2004 criteria based on persistent fever, splenomegaly, cytopenias, hypertriglyceridemia, hypofibrinogenemia, and hyperferritinemia, fulfilling six of the eight diagnostic criteria. Tropical infections were excluded with negative malaria and dengue serology and absence of other markers. She was started on hydroxychloroquine and prednisolone (40 mg), followed by pulse methylprednisolone therapy due to ongoing anemia and fever. The patient's condition improved with treatment, and a renal biopsy was planned on an outpatient basis to evaluate for lupus nephritis.

Case 3

An 18-year-old female presented with a seven-day history of fever, jaundice, abdominal pain, vomiting, and loose stools. The fever had an abrupt onset, reaching 104°F, and the abdominal pain was localized to the right upper quadrant. On examination, she was icteric, had bilateral pedal edema, and exhibited tenderness in the right upper abdomen. Initial laboratory findings revealed anemia, leukocytosis, and thrombocytopenia. Dengue infection was confirmed by positive IgM and NS1 antigen tests, while screening for other tropical infections (scrub typhus, leptospirosis, malaria, typhoid, hepatitis A and E) was negative. Despite receiving intravenous fluids, antibiotics (ceftriaxone and doxycycline), and antimalarial therapy (artesunate), the fever persisted. Elevated ferritin and triglycerides with low fibrinogen levels supported the diagnosis of secondary Hemophagocytic Lymphohistiocytosis (HLH) (Table 1). The diagnosis of secondary HLH was established according to the HLH-2004 criteria based on persistent fever, cytopenias, hypertriglyceridemia, hypofibrinogenemia, and hyperferritinemia, fulfilling five of the eight diagnostic criteria. Secondary HLH is commonly triggered by severe infections like dengue, resulting in hyperinflammation and multi-organ involvement. The patient was started on dexamethasone, and her condition stabilized with the addition of blood transfusions and iron supplementation to manage anemia. After 10 days of hospitalization, she showed clinical improvement and was discharged with follow-up plans. At discharge, she was advised to maintain adequate oral hydration, monitor blood glucose levels at home, and continue dexamethasone 4 mg twice daily for two weeks, followed by 2 mg twice daily for two weeks, and then 2 mg once daily for one week. She was also prescribed pantoprazole 40 mg once daily before meals for six weeks, paracetamol 650 mg as needed for fever, calcium 500 mg once daily for six weeks, cefixime 200 mg once daily for two days, ursodeoxycholic acid 300 mg twice daily, and rifaximin 550 mg twice daily.

Materials and methodology

This case series included three adolescent female patients diagnosed with secondary Hemophagocytic Lymphohistiocytosis (HLH) at a tertiary care hospital. Clinical history, examination findings, and laboratory investigations were recorded, and the diagnosis of secondary HLH was established according to the HLH-2004 diagnostic criteria. Bone marrow examination was performed when clinically indicated. Hemoglobin, leukocyte count, platelet count, and differential counts were analyzed using Sysmex XN 1000 (Sysmex Corporation, Japan). Absolute differential leukocyte counts were calculated by multiplying the total leukocyte count by the percentage of each cell type divided by 100. C-Reactive Protein (CRP), liver enzymes (AST, ALT), bilirubin, LDH, triglycerides, and urinary microalbumin-creatinine ratio were measured using AU680 analyzer (Beckman Coulter, USA). Ferritin, fibrinogen, and procalcitonin were assessed using ADVIA Centaur XP (Siemens Healthineers, Germany) and Vitros 5600 (Ortho Clinical Diagnostics, USA). Blood cultures were performed using BD BACTEC FX (Becton Dickinson, USA), and the Direct Coombs Test (DCT) was done using Bio-Rad kits (Bio-Rad Laboratories, USA). The Direct Coombs Test (DCT) was performed using Bio-Rad ID-Card (Bio-Rad Laboratories, USA), which detects both IgG and C3d on red blood cells. Coagulation parameters including PT/INR were measured using STA Compact Max (Diagnostica Stago, France). Ultrasound of the abdomen and chest X-ray were performed using Ysio Max (Siemens Healthineers, Germany), while CECT scans were done with Revolution CT (GE Healthcare, USA). Autoimmune profiling, including ANA patterns, was performed using Indirect Immunofluorescence Test (IIFT) on EUROIMMUN kits (EUROIMMUN, Germany).

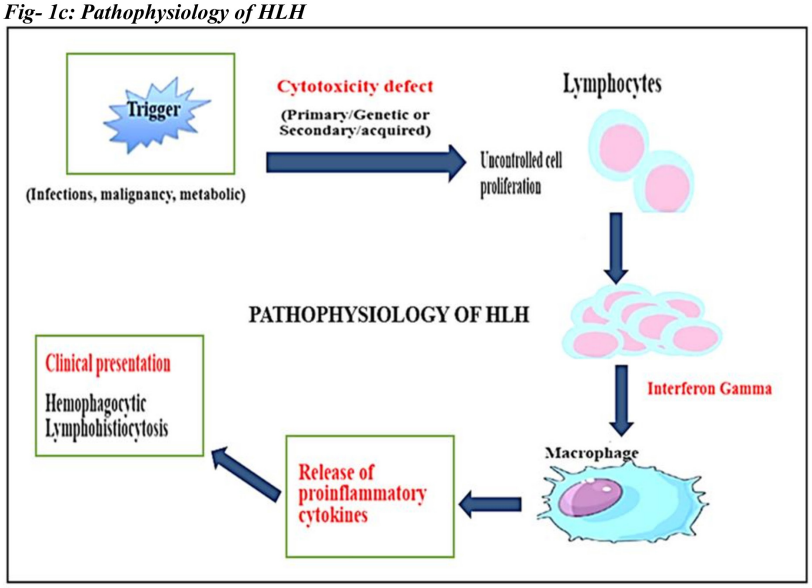
Discussion

In recent years, several innovative strategies have been developed to evaluate patients exhibiting clinical and laboratory findings suggestive of Hemophagocytic Lymphohistiocytosis (HLH). HLH is a severe hyperinflammatory condition with the potential to rapidly advance to multiple organ failure (MOF) and can be fatal [1,2]. It is categorized into primary (familial), arising from genetic defects in the perforin signaling pathway, and secondary (acquired), which occurs in the context of infections, malignancies, or autoimmune diseases [3,4]. When HLH is associated with autoimmune disorders, it is termed macrophage activation syndrome (MAS). The defining feature of HLH is unregulated immune activation involving excessive stimulation of antigen-presenting cells (macrophages and histiocytes) along with CD8+ T cells. This leads to persistent hypercytokinemia, a key driver in HLH pathogenesis, contributing to tissue injury and MOF [5]. Elevated pro-inflammatory cytokines such as TNF-α, IFN-γ, IL-6, IL-12, and IL-18 further amplify the immune response, leading to systemic inflammation (Figure 1c). Diagnosing HLH remains complex due to its non-specific clinical manifestations. The International Histiocyte Society has proposed diagnostic criteria (Table 3) requiring at

least five out of eight specified clinical and laboratory parameters, including fever, splenomegaly, cytopenias, hypertriglyceridemia or hypofibrinogenemia, elevated ferritin, raised soluble IL-2 receptor (sCD25), decreased Natural Killer (NK) cell activity, and the presence of hemophagocytosis. In secondary HLH, infections, particularly viral infections like Epstein-Barr virus (EBV), are the most common precipitating factors. Additional causes of HLH include autoimmune conditions such as systemic lupus erythematosus (SLE), juvenile idiopathic arthritis, vasculitis, and certain cancers like lymphoma and leukaemia. In HLH, excessive immune activation stems from the inability of cytotoxic CD8+ T cells and natural killer (NK) cells to clear target cells effectively, resulting in sustained immune activation and macrophage stimulation. The presence of hemophagocytosis along with elevated cytokines, particularly interferon-gamma (IFN-γ), leads to increased ferritin and TNF-α levels, contributing to key HLH manifestations such as hyperferritinemia, hypertriglyceridemia, and hepatosplenomegaly [6-9]. Recent advances in the understanding of HLH pathobiology have demonstrated that persistent activation of cytotoxic T lymphocytes and macrophages results in uncontrolled cytokine release, leading to immune-mediated tissue injury and multiorgan dysfunction. Improved understanding of these immunological pathways has facilitated the development of targeted therapeutic strategies directed against specific inflammatory mediators [13,15]. According to Ohga S et al., virus-associated hemophagocytic syndrome (VAHS), a subtype of secondary HLH, is characterized by high circulating levels of pro-inflammatory cytokines, especially IFN-γ and TNF-α during active disease, with IFN-γ serving as a critical marker of disease activity and a potent macrophage activator during hypercytokinemia [7]. Patrick DA et al. reported that secondary HLH is associated with elevated pro-inflammatory cytokines including IL-6 and IL-8, driving the hyperinflammatory state that can result in multiple organ failure (MOF), similar to what is observed in trauma patients. They also noted that a delayed increase in soluble Intercellular Adhesion Molecule-1 (sICAM-1) may indicate endothelial injury or inflammation, which is commonly seen as secondary HLH progresses [8]. Henter JI et al. described the HLH-2004 diagnostic criteria, which require fulfilling five out of eight parameters, including the recently added markers of hyperferritinemia, elevated soluble IL-2 receptor levels, and decreased or absent NK-cell activity, thereby improving diagnostic accuracy for secondary HLH (Table 2) [9]. These diagnostic markers are crucial for recognizing secondary HLH triggered by infections, autoimmune conditions, and malignancies. Although the HLH-2004 diagnostic criteria remain the cornerstone for diagnosing HLH, recent recommendations emphasise that these criteria should be interpreted alongside clinical judgement because several laboratory abnormalities may evolve during the course of illness. The North American Consortium for Histiocytosis (NACHO) recommends integrating serial laboratory investigations with clinical assessment to facilitate early diagnosis and timely initiation of treatment, particularly in critically ill patients [14].

Figure 1c: Pathophysiology of HLH.

Fig- 1c: Pathophysiology of HLH



One of the greatest challenges in diagnosing secondary HLH is its considerable clinical overlap with sepsis, severe infections, autoimmune disease flares, macrophage activation syndrome, and hematological malignancies. Hyperferritinemia, although a key diagnostic marker, is not specific to HLH and may also be observed in sepsis, acute liver injury, adult-onset Still's disease, and malignancies. Therefore, ferritin should be interpreted in conjunction with other clinical and laboratory findings. The concurrent presence of persistent fever, cytopenias, hypertriglyceridemia, hypofibrinogenemia, hepatosplenomegaly, and evidence of hemophagocytosis substantially increases the likelihood of HLH. Accordingly, application of the HLH-2004 diagnostic criteria together with careful clinical assessment is essential for differentiating HLH from these mimicking conditions.

Table 2: Diagnostic criteria for HLH-2004 requiring five of eight clinical and laboratory parameters for diagnosis.

Diagnostic criteria for haemophagocytic lymphohistiocytosis (HLH)
A. Molecular diagnosis consistent with HLH
or
B. Diagnostic criteria for HLH fulfilled (5 of the 8 criteria below)
Fever
Splenomegaly
Cytopenias (affecting ≥ 2 of 3 lineages in the peripheral blood)
Haemoglobin < 9 g/L
Neutrophils < 1.0×10^9 µ/L
Platelets < 100×10^3 µ/L
Hypertriglyceridaemia and/or hypofibrinogenaemia
Fasting triglycerides ≥ 3.0 mmol/L (i.e. ≥ 265 mg/dL)
Fibrinogen ≤ 1.5 g/L (i.e. ≤ 150 mg/dL)
Haemophagocytosis in bone marrow, spleen or lymph nodes
Low or no NK-cell activity
Ferritin ≥ 500 ng/L
sCD25 (i.e. soluble IL-2 receptor) ≥ 2400 U/mL

HLH: Hemphagocytic Lymphohistiocytosis, NK: natural killer.

Differential diagnosis

Secondary HLH shares several clinical and laboratory features with sepsis, severe systemic infections, autoimmune flare, macrophage activation syndrome (MAS), acute liver failure, thrombotic microangiopathy, and hematological malignancies. Consequently, differentiating HLH from these conditions requires careful integration of clinical presentation, laboratory findings, and the HLH-2004 diagnostic criteria.

Case 1: Persistent fever with positive blood culture initially suggested bacterial sepsis. However, persistent fever despite appropriate antibiotics, hepatosplenomegaly, cytopenias, markedly elevated ferritin, hypofibrinogenemia, and bone marrow hemophagocytosis fulfilled the HLH-2004 diagnostic criteria, supporting infection-triggered secondary HLH.

Case 2: The principal differential diagnoses included active systemic lupus erythematosus (SLE) flare, autoimmune hemolytic anemia (AIHA), macrophage activation syndrome (MAS), and infection. Although SLE accounted for the autoimmune manifestations, the coexistence of persistent fever, splenomegaly, cytopenias, Hyperferritinemia, hypofibrinogenemia, and hypertriglyceridemia was more consistent with secondary HLH occurring in association with SLE.

Case 3: Severe dengue infection alone can cause thrombocytopenia and hepatic dysfunction. However, persistent fever, marked Hyperferritinemia, hypertriglyceridemia, hypofibrinogenemia, progressive cytopenias, and multiorgan involvement strongly favoured dengue-associated secondary

HLH rather than uncomplicated severe dengue.

Although the final immunopathological pathway in HLH is characterized by excessive macrophage activation and a cytokine storm, the initiating mechanisms differ according to the underlying trigger. In Case 1, Pseudomonas aeruginosa bacteremia likely induced persistent activation of the innate immune system through pathogen-associated molecular patterns (PAMPs), resulting in sustained macrophage and cytotoxic T-cell activation with excessive cytokine release. Failure to adequately regulate this inflammatory response may have culminated in secondary HLH. In Case 2, systemic lupus erythematosus (SLE) with autoimmune hemolytic anemia (AIHA) created a state of chronic immune dysregulation characterized by persistent activation of autoreactive lymphocytes, immune complex formation, complement activation, and excessive cytokine production. These mechanisms predispose susceptible patients to macrophage activation syndrome; a form of secondary HLH associated with autoimmune diseases. In Case 3, severe dengue infection is known to trigger intense immune activation through direct viral stimulation of monocytes, macrophages, and T lymphocytes, resulting in marked production of pro-inflammatory cytokines including interferon-γ, tumour necrosis factor-α, and interleukins. This exaggerated immune response contributes to endothelial dysfunction, multiorgan involvement, and progression to dengue-associated secondary HLH. Despite the differing triggers, all three cases converged on a common pathway of uncontrolled immune activation, resulting in the characteristic clinical and laboratory manifestations of HLH [5–7,12–15].

The clinical laboratory plays a pivotal role in the early recognition of secondary HLH by identifying characteristic biochemical and hematological abnormalities that may precede definitive clinical diagnosis. In all three patients, routine laboratory investigations including complete blood counts, liver function tests, serum ferritin, fibrinogen, triglycerides, lactate dehydrogenase (LDH), and inflammatory markers provided the initial evidence of an underlying hyperinflammatory syndrome. Persistent cytopenias involving two or more cell lineages, marked hyperferritinemia, hypofibrinogenemia, hypertriglyceridemia, and elevated liver enzymes should prompt clinicians to consider HLH, particularly in patients with persistent fever and multiorgan involvement. A stepwise laboratory approach beginning with complete blood count and inflammatory markers, followed by serum ferritin, triglycerides, fibrinogen, liver function tests, and bone marrow examination when clinically indicated, can facilitate early application of the HLH-2004 diagnostic criteria. Timely communication of these critical laboratory abnormalities between the laboratory and the treating clinicians is essential for prompt diagnosis, early initiation of immunosuppressive therapy, and improved clinical outcomes.

Recent evidence further supports serial monitoring of ferritin, triglycerides, fibrinogen, complete blood counts, liver function tests, and inflammatory markers for evaluating disease progression and response to therapy. Integration of these laboratory parameters with the HLH-2004 diagnostic criteria enables earlier recognition of HLH and facilitates prompt therapeutic intervention [13,14]. Additionally, Rosenbloom et al. demonstrated that in patients with Systemic Inflammatory Response Syndrome (SIRS), increased expression of CD11b and CD35 on polymorphonuclear neutrophils (PMNs) correlated with serum IL-6 levels and the severity of organ dysfunction, suggesting that PMN activation is linked to inflammation and organ failure in these patients [10]. Haddad E et al. noted that central nervous system (CNS) involvement frequently occurs in secondary HLH, often manifesting early with symptoms such as seizures and meningitis, and that bone marrow transplantation (BMT) performed after initial remission offers the best chance for halting CNS disease progression and achieving long-term remission [11]. Furthermore, advances in the understanding of HLH immunopathogenesis have resulted in the development of targeted therapies, including interferon-γ blockade and Janus kinase (JAK) inhibitors, particularly for patients with refractory or relapsed disease. Although corticosteroids and etoposide remain the standard initial therapy, these newer treatment strategies represent promising adjuncts in selected patients [15,16].

In summary, the integration of clinical, biochemical, and genetic findings is essential in diagnosing HLH, enabling timely intervention that can greatly enhance patient prognosis.

Table 3: HLH-2004 diagnostic criteria fulfilled by each patient.

HLH-2004 Criterion	Case 1	Case 2	Case 3
Fever	✓	✓	✓
Splenomegaly	✓	✓	✗
Cytopenias	✓	✓	✓
Hypertriglyceridemia	✗	✓	✓
Hypofibrinogenemia	✓	✓	✓
Ferritin ≥500 ng/mL	✓	✓	✓
Hemophagocytosis	✓	✗ (Not performed)	✗ (Not performed)
NK-cell activity	Not done	Not done	Not done
Soluble CD25	Not done	Not done	Not done
Criteria fulfilled	**6/8**	**6/8**	**5/8**

HLH-2004 diagnostic criteria fulfilled by the three patients. Diagnosis of secondary HLH was established according to the HLH-2004 guidelines, requiring fulfilment of at least five of the eight diagnostic criteria.

Conclusion

HLH represents a heterogeneous disorder with various underlying causes and is likely underdiagnosed, contributing to its significant morbidity and mortality. This case series describes three patients with secondary HLH, each arising from different triggers: a bacterial infection (Pseudomonas aeruginosa), an autoimmune condition (SLE with AIHA), and a viral infection (severe dengue). Laboratory tests play a critical role in diagnosing secondary HLH, with findings such as hyperferritinemia, elevated liver enzymes (aspartate transaminase [AST] and alanine transaminase [ALT]), hypertriglyceridemia, and hypofibrinogenemia, alongside histopathological confirmation through evidence of hemophagocytosis. Notably, hypofibrinogenemia was identified as a marker associated with poorer outcomes. Prompt recognition and management of the underlying causes are essential to improving prognosis in secondary HLH. Nevertheless, ongoing efforts are needed to enhance disease awareness and optimize treatment approaches for affected patients.

Acknowledgments

We are thankful to the technical staff and doctors who contributed towards laboratory testing and patient management. We acknowledge the contributions of patient, their attendants, nuring officers and technical staff in providing us relevant kits to carry out the research.

Ethical clearance

According to institutional policy, ethical approval was waived by the Institutional Ethics Committee for this case series.

Written informed consent for publication was obtained from the legal guardian of the minor patient (Case 1) and from the adult patients (Cases 2 and 3) prior to inclusion in the study.
Ethical clearance is waived off for case reports by IEC, AIIMS Patna.

Financial support

None.

Conflict of interest

Authors declare no conflict of interest.

Author contributions

1. LPK, MM: conception or design of the work, acquisition, analysis, and interpretation of data for the work
2. LPK, SK, MM: analysis and interpretation of data,
3. LPK, MM: Drafting the work or revising it critically for important intellectual content
4. SK,RK,JP: providing clinically relevant samples for assay
5.RK,JP:Clinical management

All authors have approved of the final draft for publication purpose and are accountable for the accuracy and integrity of the work.

Supplementary material

No.
Informed consent was taken from patient concerned for publishing data and case details in a medical journal.

Data availability statement

All relevant data have been provided along with the manuscript.

References

1. Gauvin F, Toledano B, Champagne J, Lacroix J: Reactive hemophagocytic syndrome presenting as a component of multiple organ dysfunction syndrome. Crit Care Med 2000, 28:3341-3345.
2. Karapsınar B, Yilmaz D, Balkan C, Akin M, Ay Y, Kavakli K: An unusual cause of multiple organ dysfunction syndrome in the pediatric intensive care unit: hemophagocytic lymphohistiocytosis. Pediatr Crit Care 2009, 10:285-290.
3. Henter JI, Elinder G, Ost A: Diagnostic guidelines for Hemophagocytic lymphohistiocytosis. The FHL Study Group of the Histiocyte Society. Semin Oncol 1991, 18:29-33.
4. Fujiwara F, Hibi S, Imashhuku S: Hypercytokinemia in hemophagocytic syndrome. Am J Pediatr Hematol Oncol 1993, 15:92-98.
5. Janka GE: Hemophagocytic syndromes. Blood Rev 2007, 21:245-253.
6. Imashuku S, Hibi S: Cytokines in Hemophagocytic syndrome. Br J Haematol 1991, 77:438-440.
7. Ohga S, Matsuzaki A, Nishizaki M, Nagashima T, Kai T, Suda M, Ueda K: Inflammatory cytokines in virus-associated hemophagocytic syndrome Interferon gamma as a sensitive indicator of disease activity. Am J Pediatr Hematol Oncol 1993, 15:291-298.
8. Patrick DA, Moore FA, Moore EE, Biff WL, Sauaia A, Barnett CC Jr: The inflammatory profile of interleukin-6, interleukin-8, and soluble intercellular adhesion molecule-1 in postinjury multiple organ failure. Am J Surg 1996, 172:425-429.
9. Henter JI, Horne A, Arico M, Egeler RM, Filipovich AH, ImashukuS, et al. HLH-2004: diagnostic and therapeutic guidelines for hemophagocytic Lymphohistiocytosis. Pediatr Blood Cancer.2007;48:124-31.
10. Rosenbloom AJ, Pinsky MR, Bryant JL, Shin A, Tran T, Whiteside T: Leukocyte activation in the peripheral blood of the patients with cirrhosis of the liver and SIRS. Correlation with serum interleukin-6 levels and organ dysfunction. JAMA 1995, 274:58-65.
11. Haddad E, Sulis M-L, Jabado N, Blanche S, Fischer A, Tardieu M. Frequency and severity of central nervous system lesions in hemophagocytic lymphohistiocytosis. Blood. 1997; 89:794–800.
12. Canna SW, Marsh RA. Paediatric hemophagocytic lymphohistiocytosis. Blood. 2020;135(16):1332–1343. Doi: 10.1182/blood.2019000936.
13. Bode SFN, Lehmberg K, Maul-Pavicic A, Vraetz T, Janka G, zur Stadt U, et al. Recent advances in the diagnosis and treatment of hemophagocytic lymphohistiocytosis. Arthritis Research & Therapy. 2012;14(3):213.
14. Jordan MB, Allen CE, Greenberg J, Henry M, Hermiston ML, Kumar A, et al. Challenges in the diagnosis of hemophagocytic lymphohistiocytosis: Recommendations from the North American Consortium for Histiocytosis (NACHO). Paediatric Blood & Cancer. 2019;66: e27929.
15. Al-Samkari H, Berliner N. Hemophagocytic lymphohistiocytosis. Annual Review of Pathology. 2018; 13:27-49.
16. Jordan MB, Allen CE, Weitzman S, Filipovich AH, McClain KL. How I treat hemophagocytic lymphohistiocytosis. Blood. 2011;118(15):4041-4052.

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.

Commentary

Medical Laboratory Accreditation Programs: A Comparative Review of International and National Frameworks

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

[1]*Productivity and Quality Institute (PQI), Arab Academy of Science, Technology, and Maritime Transport (AASTMT), Cairo, Egypt*
[2]*Department of Basic and Applied Science, College of Engineering and Technology, Productivity and Quality Institute (PQI), Arab Academy for Science, Technology and Maritime Transport (AASTMT), Cairo, Egypt*
[3]*Center for Informatics Science (CIS), School of Information Technology and Computer Science, 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
E-mail: masaid@nu.edu.eg
Phone: +201001662403

Keywords

Medical Laboratory Accreditation, CAP LAP, ISO 15189, CAP-ISO 15189, JCI Accreditation, National Accreditation Systems, Patient Safety, Quality Management Systems

Abstract

Clinical laboratories play a critical role in modern healthcare, supporting nearly 70% of clinical decision-making related to diagnosis, monitoring, and prognosis. However, errors across the pre-analytical, analytical, and post-analytical phases can significantly compromise patient safety. Laboratory accreditation has emerged as a key strategy to enhance quality, reduce errors, and strengthen healthcare systems. This narrative review critically analyzes the impact of major international and national laboratory accreditation programs, including the College of American Pathologists Laboratory Accreditation Program (CAP LAP), ISO 15189, Joint Commission International (JCI), CAP- ISO 15189, as well as national frameworks such as CBAHI, GAHAR, and NABL.A structured narrative review was conducted to synthesize available evidence on medical laboratory accreditation frameworks and their impact on quality and patient safety. The findings indicate that accreditation is consistently associated with improvements in quality management systems, staff competency, process standardization, and patient safety. Comparative analysis suggests that laboratory-specific accreditation models, such as CAP and ISO 15189-based frameworks, primarily enhance technical and operational performance, whereas broader healthcare accreditation systems, such as JCI, facilitate integration with institutional patient safety and governance structures. National accreditation programs play a critical role in adapting international standards to local regulatory and healthcare contexts. In conclusion, laboratory accreditation represents a fundamental mechanism for improving healthcare quality and patient safety. Integrating multiple accreditation frameworks, particularly CAP LAP with ISO 15189, may provide a comprehensive approach to sustainable quality improvement, although further research using standardized performance indicators is required.

Introduction

Approximately 70% of medical decisions in the diagnosis, monitoring, and prognosis department are made using clinical laboratories [1]. For patient safety, timeliness, accuracy, and dependability are crucial. But mistakes can happen in the pre-analytical, analytical, and post-analytical stages, which can result in incorrect diagnoses, postponed therapy, or worse outcomes [2,3]. Laboratory accreditation has become a significant technique for increasing quality management systems, and regulatory compliance in medical laboratories. Independent organizations carry out accreditation to confirm adherence to certain competency and quality requirements, supporting staff competency, continuous improvement, and quality management systems (QMS) [3,10]. There are several national and international accreditation programs, including as the Joint Commission International (JCI) accreditation, ISO 15189, the hybrid CAP-ISO 15189 model, and the College of American Pathologists Laboratory Accreditation Program (CAP LAP). Concurrently, national initiatives like GAHAR (Egypt), CBAHI (Saudi Arabia) and NABL (India), GAHAR (Egypt), and CBAHI (Saudi Arabia) modify international standards to fit regional healthcare and regulatory environments. Evidence about the efficacy of accreditation is still inconsistent despite its widespread use. Many studies lack comparative analysis across frameworks and instead concentrate on individual systems or locations. Final conclusions are limited by variations in study design, outcome measures, and the use of self-reported data [2,10]. The aim of this review is to present a thorough comparative analysis of the main national and international laboratory accrediting frameworks, Exploring the utilization, conception and the effect of quality management system work processes and laboratory quality indicators. In contrast to previous studies that concentrated on individual systems, this analysis directly links accrediting components to patient safety results through a conceptual framework and includes a structured comparison of CAP LAP, ISO 15189, JCI, the CAP–ISO 15189 model and national frameworks (GAHAR, CBAHI& NABL)

Gap Analysis in Laboratory Accreditation Literature

There are still large gaps in research, practice, and regional application despite the extensive use of laboratory accreditation procedures. These inadequacies make it more difficult to assess accreditation systems' efficacy completely and prevent accreditation requirements from being translated into long-term gains in laboratory performance and patient safety.

Research Gap

The absence of standardized key performance indicators (KPIs) to impartially evaluate the effect of accreditation on laboratory performance and patient outcomes is a significant drawback of the existing literature. Most research uses self-reported data and cross-sectional designs, which limit causal inference and weaken methodological robustness [2,10]. Furthermore, there are very few longitudinal studies assessing the long-term sustainability of accreditation outcomes, especially in quickly changing healthcare systems. The necessity for internationally standardized evaluation frameworks connected to patient safety indicators is highlighted by the fact that variations in study designs and measurement instruments further diminish comparability across accreditation models.

Practice Gap

At the operational level, accreditation implementation frequently prioritizes regulatory compliance and documentation over quantifiable clinical outcomes. Excessive administrative work may lower staff engagement and divert attention from patient-centered changes, even if documentation is crucial for quality assurance [11,13]. Furthermore, as performance gains are occasionally motivated by inspection preparedness rather than incorporated continuous quality improvement systems, the longevity of quality improvements following certification cycles is still an issue [20].

Regional Gaps

Regional disparities continue to be a major problem, especially in the Middle East and low- and middle-income nations. Even though national accreditation frameworks like CBAHI and GAHAR have improved regulatory standards, there is still a dearth of published data assessing their long-term effects on patient outcomes [26, 27]. Variability in implementation and results between areas is further influenced by differences in financial resources, personnel expertise, and infrastructure [28].

Synthesis of Gap Analysis

Overall, the discrepancies found show that accrediting frameworks and outcome-based healthcare quality assessment are consistently out of synch. Although certification programs successfully standardize laboratory procedures, there is still limited evidence demonstrating long term improvements in the measurable patient outcome. Standardized KPIs must be developed, longitudinal study designs must be used more frequently, and clinical outcome indicators and accreditation criteria must be more closely aligned to close these gaps.

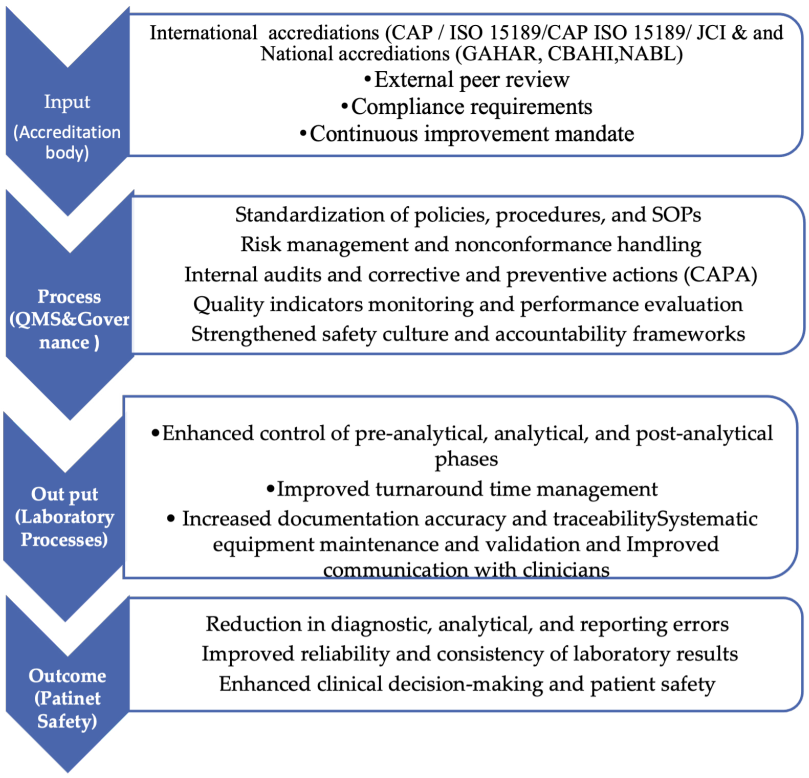
Table 1: Summary of Identified Gaps in Laboratory Accreditation Evidence.

Gap Category	Description	Implications for Knowledge & Practice
Research Gaps	Lack of established Key Performance Indicators (KPIs) to evaluate how certification affects patient outcomes and laboratory performance.	Meta-analysis is hampered by limitations on objective evaluation and cross-study comparability.
	Cross-sectional approaches and self-reported data predominate; longitudinal research evaluating the sustainability of gains is scarce.	Limits the ability to draw conclusions about causality; it is uncertain whether the advantages of accreditation are temporary or long-lasting.
	Differences across accreditation models in terms of measurement methods, study designs, and outcome definitions.	Decreases comparability and emphasizes the necessity of evaluating frameworks that are universally standardized.
Practice Gaps	Prioritizing regulatory compliance and documentation over quantifiable clinical results or a continuous quality improvement culture.	Increases administrative work, which could demotivate employees and divert attention from patient-centered treatment.
	Poor quality improvement sustainability after accreditation; changes are frequently motivated more by preparedness for inspections than by integrated quality systems.	Following accreditation cycles, performance may deteriorate; there is a disconnect between outcome-based and compliance-based quality management.
Regional Gaps	Despite the existence of national frameworks (such as CBAHI, GAHAR, and NABL), there is little published data on the long-term effects of accreditation in low- and middle-income nations in the Middle East.	In many areas, policymaking and successful implementation are hampered by a lack of context-specific evidence.
	International standards are not consistently implemented or adapted due to differences in infrastructure, worker competency, and financial resources.	Impacts the uniformity of accrediting results and raises the possibility of uneven quality gains in different areas.

Conceptual Framework of Laboratory Accreditation and Patient Safety

In order to clarify the ways in which laboratory accreditation enhances patient safety and healthcare quality, this review article offers a conceptual framework. The framework provides a methodical depiction of how accreditation results in quantifiable gains in laboratory performance and clinical outcomes by integrating structural, process, and outcome characteristics. At the structural input level, accreditation frameworks such as CAP LAP, ISO 15189, JCI, and national systems (e.g., CBAHI, GAHAR, NABL) create uniform standards for technical proficiency, laboratory governance, and quality management systems (QMS). These components serve as the cornerstone for guaranteeing adherence to global norms and industry best practices. At the process level, by implementing standard operating procedures (SOPs), internal audits, proficiency testing (PT/EQA), and ongoing staff competency assessments, laboratories operationalize accreditation standards. Throughout the pre-analytical, analytical, and post-analytical stages of testing, these procedures seek to improve uniformity and lessen variability. At the output level, accreditation is associated with improvement in quality processes, laboratory performance and compliance with standardized procedures are linked to accreditation. These enhancements show improved laboratory service dependability and operational efficiency. At the outcome level, these changes are intended safer clinical decision making and improved quality of care, however, direct evidence linking the accreditation to measurable patient outcomes remain limited in the current literature. Furthermore, accreditation encourages a culture of accountability and ongoing quality improvement in healthcare institutions. This framework extended beyond descriptive models by using intermediary process and performance indicators to directly connect accrediting frameworks to clinical results. It fills a significant gap in the existing literature by highlighting the significance of matching accrediting requirements with quantifiable patient-centered outcomes.

Figure 1: Conceptual framework illustrating the relationship between laboratory accreditation structures, quality processes, operational outputs, and patient safety outcomes.



Accreditation Processes Frameworks

Laboratory accreditation are organized procedures intended to guarantee patient safety, technical proficiency, and quality control in diagnostic services. Although accrediting systems have similar goals, they differ in terms of structure, regulatory philosophy, and evaluation scope. A comparative synthesis of the main national and international accrediting process types is presented in this section.

CAP Laboratory Accreditation Program (LAP)

The highly specialized, peer-based College of American Pathologists (CAP) Laboratory Accreditation Program emphasizes technical proficiency, analytical quality, and ongoing laboratory improvement. The foundation of the accreditation cycle is continuous compliance monitoring, discipline-specific assessment instruments, and external peer review. The use of comprehensive, department-specific checklists that cover all stages of laboratory testing, including pre-analytical, analytical, and post-analytical procedures, is a distinguishing characteristic of CAP. Technical rigor and benchmarking validity are improved because on-site inspections are carried out by working laboratory personnel. To ensure ongoing compliance as opposed to one-time certification, the program uses a continuous improvement cycle that includes application review, tailored checklist assignment, peer inspection, deficiency remedy, and periodic evaluation.

Joint Commission International (JCI)

JCI accreditation follows a hospital-centered model in which laboratory services are evaluated as part of an integrated healthcare system. Instead of focusing on intricate analytical laboratory methods, the strategy prioritizes corporate governance, leadership, and patient safety culture. The tracer technique, which assesses patient care paths across departments and connects laboratory results to clinical decision-making processes, is a crucial methodological component. This method enables evaluation of system-level safety performance, communication efficacy, and workflow integration. In general, JCI prioritizes patient safety alignment and institutional quality over laboratory-specific technical depth.

ISO 15189 Accreditation Framework

An internationally accepted standard that outlines standards for competence and quality in medical laboratories is ISO 15189. It is based on a risk-oriented quality management system (QMS) approach, in contrast to checklist-driven methods. The framework highlights: Standardization of laboratory procedures, Personnel competency, both external quality

evaluation and internal quality control, risk management and ongoing improvement. While preserving worldwide uniformity in laboratory quality standards, ISO 15189 offers implementation flexibility that permits adaption across various healthcare systems.

CAP ISO 15189 Hybrid Accreditation Model

This model is an integrated strategy that blends the College of American Pathologists' (CAP) peer-review inspection methodology with the structured quality management system (QMS) standards of ISO 15189. The goal of this model is to close the gap between technical laboratory evaluation and system-based standardization. It integrates CAP's comprehensive discipline-specific checklists and on-site peer review processes, in contrast to standard ISO 15189 accreditation, which is mostly systems-oriented.

National Accreditation Systems (CBAHI, GAHAR, NABL)

To ensure regulatory alignment and implementation feasibility, national accreditation frameworks are created to localize international standards inside particular healthcare systems. Patient safety, clinical governance, and quality system standardization are the main areas of concentration for the Central Board for Accreditation of Healthcare Institutions (CBAHI) in Saudi Arabia. To improve national healthcare quality and system accountability, Egypt's General Authority for Healthcare Accreditation and Regulation (GAHAR) offer a legally regulated accreditation framework. With a focus on technical proficiency and metrological traceability, the National Accreditation Board for Testing and Calibration Laboratories (NABL) in India enforces ISO 15189-based standards. Together, these frameworks show how national healthcare objectives and resource circumstances have been adapted to global standards.

Comparative Synthesis of Accreditation Process Models

All accreditation models share fundamental concepts, such as external evaluation, process standardization, and ongoing quality improvement, despite variations in structure and application. But their emphasis is very different. Technical correctness, analytical dependability, and internal quality processes are the main focus of laboratory-specific models like CAP and ISO 15189. On the other hand, institutional models like JCI incorporate laboratory services into larger frameworks for patient safety and healthcare governance. National certification systems provide contextual relevance and regulatory compliance by bridging local healthcare rules with international standards. CAP–ISO 15189 integration, combine the methodological rigor of peer-based inspection with the structural flexibility of ISO standards, offering a more comprehensive pathway for laboratory quality assurance. Overall, accreditation frameworks should be understood as complementary rather than competing systems, each contributing differently to laboratory quality improvement and patient safety enhancement.

Methods

Study Design

The purpose of this narrative review article is intended to compile and evaluate the data that is currently available about medical laboratory accreditation frameworks and their effects on patient safety and quality management systems. This approach, in contrast to systematic reviews, enables a thorough and interpretive assessment of a variety of sources, such as institutional reports, guidelines, and empirical investigations.

Search Strategy and Data Sources

For the years 1992–2025, a thorough search of the literature was done in PubMed, ScienceDirect, BMC Health Services Research, Scopus, and Google Scholar. WHO guidelines and the official websites of accrediting organizations (CAP, JCI, NABL, CBAHI, and GAHAR) were also carefully searched.

Study Selection

A structured literature search was conducted across major search databases, including Scopus, PubMed, ScienceDirect, and Google Scholar, covering publications from 1992 to 2023. Additional sources included official guidelines and documents from accreditation bodies such as CAP, ISO15189, JCI, CAP-ISO 15189, CBAHI, GAHAR, and NABL During the initial literature search, 67 papers were found and downloaded. Relevance to medical laboratory accreditation was evaluated by screening titles and abstracts. finally, 35 sources were included after applying relevance, quality and scope of the review criteria. The priority was given to the major accreditation standards, official guidelines, highly cited empirical studies and publications directly addressing the laboratory accreditation programs and patient safety outcomes.28 peer-reviewed original articles or systematic reviews, and 7 official guidelines/ reports (CAP, JCI, WHO, CBAHI, GAHAR, NABL) [1–35].

Selection consideration

Inclusion Criteria: consisted of: Peer-reviewed articles, reports, and official publications, studies describing accreditation processes, methodologies, or outcomes, studies evaluating the impact of accreditation on laboratory quality, patient safety, or operational efficiency, publications written in English. Exclusion Criteria: The exclusion criteria included: Studies focused exclusively on analytical techniques without relevance to accreditation systems, articles that did not meaningfully address accreditation frameworks or their impact a non-English publication.

Data Extraction and Synthesis

A descriptive summary of pertinent data from the chosen research was evaluated. Key themes, including as operational performance, patient safety, and quality management systems, were found and categorized into primary domains. The results

from several certification modes were integrated using a thematic narrative synthesis approach.

Methodological Considerations

This study is categorized as a narrative review and does not follow a formal systematic review protocol, and no risk-of-bias assessment was performed.

Data Presentation Approach

To make it easier to compare and evaluate the results across accreditation methods, the synthesis findings were arranged into thematic categories. To emphasize variations in scope, technique, and influence on laboratory and healthcare quality systems, a comparative study of accrediting systems was conducted.

Summary of Methodological Approach

This narrative study offers a thorough comparative examination of the main laboratory and healthcare certification schemes by combining data from reputable accreditation sources and peer-reviewed literature. Despite the non-systematic nature of the review technique, the structured approach used guarantees transparency and thematic consistency in synthesis.

Review Protocol

Because this review was not systematic, it did not adhere to a registered protocol (such as PROSPERO) however, the search and selection process was documented transparently as described above.

Results

The summary of results from 35 included research shows that laboratory accreditation has a consistent but varied effect on operational performance, quality systems, and patient safety-related procedures. Based on recurrent domains found in the literature, the results are presented topically.

Impact on Quality Management Systems (QMS) and Operational Efficiency

Improvements in quality management system implementation and laboratory operational procedures are linked to laboratory accreditation in all of the included studies. These include strengthened internal audit systems in authorized laboratories, better adherence to standard operating procedures (SOPs), and improved documentation standards [2,10].

Laboratory-Specific Programs

Improvements in adherence to discipline-specific requirements are reported in studies assessing the CAP Laboratory Accreditation Program (CAP LAP), which are bolstered by standardized checklists and organized peer-review inspection tools [12,13]. In a similar vein, research on ISO 15189-accredited laboratories - including those in the NABL framework - reports improvements in process standardization, risk management system deployment, and organized methods for ongoing quality improvement [28]. Improvements in staff training initiatives, process standardization, and equipment maintenance systems are among the reported operational advancements [2].

Institution-Integrated Programs

Findings from studies on Joint Commission International (JCI) show that the employment of tracer techniques for cross-departmental process evaluation and standardized workflows is linked to laboratory services incorporated into hospital-wide accrediting systems [10,25].

National Frameworks

Studies analyzing the CBAHI program in Saudi Arabia report improved institutional application of safety-related regulations and more uniformity of laboratory operations [26]. In studies addressing the GAHAR framework (Egypt), Results show increased use of standardized laboratory methods and progressive harmonization with national healthcare governance systems [27].

Impact on Patient Safety Related processes and Quality of Care

Most included studies investigated surrogate measures of patient safety rather than direct clinical outcomes. Reported gains generally addressed laboratory performance parameters such as error reduction, turnaround time, process standardization, and quality assessment performance.

Laboratory performance outcomes

Studies report decreases in pre-analytical and post-analytical mistakes, especially in specimen handling, identification, and reporting procedures, in labs accredited under CAP and ISO 15189 frameworks [2,3]. Laboratories accredited under CAP and NABL systems have also reported improved performance in external quality assessment (EQA/PT schemes), which reflects improved technical performance and analytical consistency [12,28].

Healthcare service-related outcomes

Studies involving JCI-accredited institutions and national accreditation systems such as CBAHI report better collaboration between clinical and laboratory services, especially in the areas of ordering tests, communicating results, and reporting procedures [10,26]. Improved turnaround time consistency and more consistent reporting procedures across healthcare settings are also linked to standardization of laboratory workflows [18,22].

Professional Attitudes Toward Accreditation

Studies evaluating staff perceptions report opinions on laboratory accrediting schemes are typically favorable. Improved organizational structure, more precise process

description, and more professional recognition are all linked to accreditation [11,26].Studies do, however, also note difficulties with increased documentation requirements, administrative burden, and accreditation inspection preparation [11,22]. According to research done in Saudi Arabia, laboratory staff members express worries about workload intensity and maintaining compliance between inspection cycles, but they also indicate good opinions of safety improvements linked to CBAHI accreditation [20].

Comparative Analysis Derived from the Literature

The literature shows that the features and stated results of various accrediting methods vary depending on their implementation focus and breadth.

Laboratory-Specific Programs (CAP, ISO 15189, NABL):

These models place a strong emphasis on technical proficiency, analytical quality, and process standardization at the laboratory level. Strong alignment with internal quality control systems and improvements in technical performance are regularly reported in studies [23, 24, 28].

Institutional Accreditation Programs (JCI, CBAHI, GAHAR):

The integration of laboratory services into larger hospital-wide patient safety and governance systems is the main goal of these frameworks. Results show cross-departmental coordination and organized alignment with organizational quality systems [25–27].

Hybrid Models (CAP ISO 15189): CAP accreditation embraces ISO 15189 quality management concepts while keeping a checklist-based inspection approach. Results show that comprehensive technical assessment systems and structured QMS criteria are aligned simultaneously [23].

Synthesis of Reported Findings

Improvements in quality management systems, laboratory procedures, and adherence to established standards are consistently linked to laboratory certification in all of the included research. However, variations in study design, healthcare settings, and performance metrics between studies are responsible for the observed diversity in reported outcomes. This reflects the variation in evaluation techniques seen in the literature and restricts direct comparability between accrediting regimes.

Table 1: Comparison of Key Laboratory Accreditation Models.

Feature	ISO 15189	CAP Accreditation	CAP ISO 15189	JCI Accreditation	NABL (India)	CBAHI (Saudi Arabia)	GAHAR (Egypt)
Primary Philosophy	Process competence & continual improvement	Compliance & peer benchmarking	Convergence of ISO process focus with CAP's peer-review administration	Integrated hospital-wide patient safety, and laboratories	National implementation of ISO 15189	Mandatory national healthcare quality improvement	Mandatory national healthcare regulation & quality improvement
Scope of Accreditation	Laboratories	Laboratories	Laboratories	Hospitals and laboratories	Medical laboratories, testing laboratories and calibration laboratories	Hospitals and laboratories	Hospitals and laboratories
Standard Format	High-level principles & requirements, laboratory	Detailed, prescriptive checklists	ISO 15189 standard, interpreted and assessed within CAP's framework	Comprehensive hospital standards & laboratories	ISO 15189 standard	Integrated National Standards	Integrated national standards
Inspection Process	Assessors trained in ISO 15189 standards	Peer-based inspection by laboratory professionals (working employees)	CAP-trained peer assessors conducting an ISO 15189-based assessment	Multidisciplinary survey teams	Assessments by NABL-trained assessors	Surveys by CBAHI-trained surveyors	Surveys by GAHAR-trained assessors
Validity of certificate	3 years	2 years	3 years	3 years	4 years	3 years	3 years
Standards basis	ISO 15189 Standards	CAP Checklist with detailed standards for each department	CAP & ISO 15189 Standards	JCI standards for laboratories	Based on ISO 15189 Standards	CBAHI standards (national standards)	GAHAR standards (national standards)

Table 2: Strengths and Limitations of Accreditation Models.

Accreditation Model	Scope	Strengths	Limitations	Technical Depth	Healthcare Integration	Resource Demand
CAP LAP (College of American Pathologists Laboratory Accreditation Program)	Laboratory-specific, highly specialized	Extensive discipline-specific checklists; a strong peer-review inspection system; ongoing modifications based on evidence; and a strong emphasis on improving analytical quality	Exorbitant expense, extensive planning, reliance on knowledgeable inspectors, and administrative load	Very High	Moderate	High
CAP–ISO 15189 (CAP assessment model aligned with ISO 15189 requirements)	Integrated laboratory accreditation using the ISO 15189 QMS framework and CAP inspection methodology	Integrates ISO 15189's QMS and process approach with CAP's technical rigor and peer review; enables dual compliance and international recognition; and promotes a sustainable quality culture.	Dual standard alignment makes implementation difficult; documentation must be harmonized, assessors must be cross-trained, and administrative overhead may increase.	Very High	High	Very High
ISO 15189 (International Organization for Standardization)	Global laboratory quality standard applicable across healthcare systems	International recognition, adaptability to different laboratory settings, flexibility in implementation, and a strong focus on QMS and risk management	Requires interpretation throughout implementation; compliance depth varies; discipline-specific granularity may be lacking.	High	Moderate	Moderate
JCI (Joint Commission International)	Hospital-wide healthcare accreditation system (includes laboratory standards)	Tracer approach improves process mapping; strong integration with institutional patient safety and governance; and promotion of an organization-wide quality culture	In comparison to specialized lab certification models, there is less technical depth unique to laboratories; a wider reach could lessen attention to technical details.	Moderate	Very High	High
CBAHI (Central Board for Accreditation of Healthcare Institutions, Saudi Arabia)	National healthcare system (Saudi Arabia)	Strong adherence to national healthcare principles; improves regulatory compliance and uniformity; required for numerous facilities	Institutional differences in implementation maturity; proof of results is still developing	Moderate	High	Moderate
GAHAR (General Authority for Healthcare Accreditation and Regulation, Egypt)	National healthcare accreditation framework (Egypt)	Legally mandated national standards and a structured framework for healthcare reform and system-wide quality improvement	Early implementation, scant evidence of long-term results, and sector-specific resource variability	Moderate	High	Moderate
NABL (National Accreditation Board for Testing and Calibration Laboratories, India)	Laboratory and calibration focused (ISO 15189-based)	Strong focus on technical proficiency and metrological traceability; ISO-aligned organization; supports national quality infrastructure	Limited resources, inconsistent adoption among labs, and a significant amount of paperwork for smaller labs	High	Moderate	Moderate–High

Comparative Advantages and Limitations of Accreditation Models

Laboratory accreditation systems differ substantially in their conceptual scope, regulatory objectives, and implementation methodologies. While some frameworks place a higher priority on analytical validity and technical laboratory quality, others place more emphasis on organizational governance, patient safety integration, or national healthcare alignment. These discrepancies, which range from highly specialized laboratory focused models to broader healthcare system wide approaches, reflect disparities in accreditation philosophies. In order to critically evaluate each model's relative strengths, limitations, and applicability in enhancing laboratory quality and patient safety outcomes, a systematic comparative evaluation is necessary. The following table summarizes the advantages and disadvantages of accreditation models.

Overall, the comparison analysis shows that no accreditation methodology is consistently better in every aspect of performance. The JCI model shows the best interaction with hospital-wide patient safety systems, while the CAP LAP framework shows the highest degree of technical rigor and analytical specificity. While national accreditation systems like CBAHI and GAHAR stress contextual congruence with local healthcare priorities, ISO 15189 offers a globally standardized and adaptable framework for quality management and the (CAP–ISO 15189) accreditation pathway. This dual-alignment approach promotes international recognition and thorough quality assurance while improving laboratory standards. However, because two different certification principles must be harmonized, it also adds to the complexity of implementation.

Discussion

This analysis offers a methodical synthesis of the available data on laboratory accrediting systems, emphasizing their influence on patient safety related procedures, operational performance, and quality management. The results show that accreditation is a multifaceted framework whose effects vary according on the model, scope, and healthcare setting rather than a consistent intervention.

Interpretation of Impact on Quality Management Systems

The review's findings verify that advances in operational standardization and quality management systems (QMS) are regularly linked to laboratory certification. Internal audit system implementation, standard operating procedure adherence, and structured documentation were commonly reported across several certification types. Because of their unique requirements and discipline-specific focus, laboratory-specific certification systems like CAP LAP and ISO 15189 seem to offer more technical uniformity and process control. On the other hand, by integrating quality systems into larger organizational structures, institution-based accrediting schemes like JCI and CBAHI take these advancements outside of the

lab. These results imply that while all certification models increase QMSs, the extent and depth of the impact vary according on whether the accreditation is institution-wide or laboratory-focused.

Interpretation of Patient Safety and Quality Outcomes

Based on the review findings, it appears that the relationship between accreditation and patient safety is mainly indirect and process based. Surrogate indicators of patient safety, such as reduced pre-analytical and post-Analytical errors, faster turnaround times, better compliance with standard operating procedures and better performance in external quality assessment programs, were used by most included studies. Safe clinical practice is often influenced by these measures, which reflect improvements in laboratory quality and operational performance. There are comparatively few studies that specifically examined patient-centered clinical outcomes, including delays in critical value reporting for physicians, diagnostic error rates, transfusion-related errors, and other quantifiable but not necessarily measurable errors in patient safety indicators. Laboratory process and quality indicators are the primary sources of evidence, with little direct clinical benefit. Consequently, the current literature suggests that accreditation is advantageous for patient safety by improving laboratory quality processes and minimizing diagnostic error.

Professional and Organizational Implications

The results pertaining to professional attitudes show that accreditation is typically seen favorably, especially with regard to role clarity, organizational structure, and professional credibility. Simultaneously, the adoption of certification may present operational issues due to the increased administrative workload and documentation requirements documented in multiple studies. The balance between resource burden and quality improvement is reflected in this dual perspective, especially in environments with constrained infrastructure or personnel capability. Maintaining accreditation benefits seems to depend on internal organizational commitment and a continuous quality culture in addition to standard compliance.

Comparative Interpretation of Accreditation Models

The comparative analysis highlights fundamental differences between accreditation models in terms of technical depth, system integration, and implementation approach. In terms of technical quality, analytical accuracy, and internal laboratory procedures, laboratory-focused models (CAP, ISO 15189, NABL) exhibit excellent performance. Interdisciplinary collaboration, patient safety mechanisms, and integration with hospital-wide governance are highlighted by institutional models (JCI, CBAHI, GAHAR) and Hybrid models (CAP–ISO 15189). These results suggest that certification models should be seen as complementing frameworks that address various aspects of healthcare quality rather than as directly I terchangeable.

Global Standardization versus Local Adaptation

The harmony between local contextualization and worldwide standardization is a major subject that emerged from this review. Although they offer a widely accepted framework for laboratory quality, international standards like ISO 15189 necessitate interpretation during implementation. National accreditation programs like CBAHI and GAHAR, on the other hand, modify these guidelines to take into account regional healthcare priorities, legal requirements, and resource limitations. The tendency of combining global standards with context-specific implementation strategies is exemplified by hybrid approaches, such CAP–ISO 15189. This illustrates the continuous transition in certification procedures toward adaptable, context-aware quality systems.

Implications for Practice and Future Research

The review's findings imply that laboratory certification enhances organizational performance, operational consistency, and quality systems. However, it is difficult to directly compare accrediting models due to differences in study design, outcome measures, and reporting methods. Future studies ought to concentrate on: Creating uniform performance metrics, setting up impartial patient safety measures, performing long-term research to evaluate the sustainability of accreditation results, Assessing the financial viability of various accrediting models. Financial and resource requirements should also be considered when adopting an accreditation approach. As illustrated in Table 2, accreditation systems differ greatly in implementation difficulty, staff requirements, training demands, documentation burden, and the inspection expenses. CAP and CAP–ISO 15189 pathways generally require greater organizational investment because of their extensive technical requirements and ongoing compliance activities, whereas ISO 15189 and national accreditation frameworks may offer more flexible implementation approaches depending on local regulatory and healthcare contexts. These resource considerations may influence both accreditation selection and sustainability, particularly in resource-constrained situations.

Limitations of the Review

This review has several limitations. First, it included only literature published in English. Second, the variability in study designs, outcome measures, and reporting methods limits the ability to draw definitive conclusions. Third, reliance on self-reported data in some studies introduces potential bias. Future research should focus on standardized performance indicators, objective patient safety metrics, and longitudinal studies to better quantify the impact of accreditation on laboratory and healthcare performance.

Conclusion

Across a range of healthcare settings, advancements in quality management systems, operational procedures, and laboratory

performance are regularly linked to laboratory accreditation. The review's conclusions show that certification strengthens quality processes, improves organizational structure in laboratories, and increases uniformity. However, depending on the model used, the extent of implementation, and the healthcare setting, accreditation has different effects on laboratory performance indicators and laboratory processes. While institutional and national accreditation systems like JCI, CBAHI, and GAHAR enable integration with more comprehensive patient safety and healthcare governance structures, laboratory-focused accreditation frameworks like CAP and ISO 15189 mainly support technical quality and process control. There isn't a single accrediting model that consistently outperforms others in every aspect of performance. Rather, congruence with institutional goals, legal criteria, and available resources determines how effective certification is. Combining organized quality management with in-depth technical evaluation can be accomplished through hybrid techniques, especially those that combine ISO 15189 with CAP-based assessment methodologies. To enable more reliable comparison of certification outcomes, future research should concentrate on creating objective patient safety criteria and standardized performance indicators. To assess the durability and long-term effects of accreditation on laboratory performance and healthcare quality, longitudinal studies are also required. While accreditation is frequently related with improvements in laboratory quality indicators, actual evidence indicating improvements in patient centered clinical outcomes remains limited. All things considered, laboratory certification continues to be a crucial part of quality assurance frameworks, encouraging uniformity in laboratory practice across various healthcare systems and supporting ongoing progress.

Patents

The authors declare that no patents are associated with this work.

Supplementary Materials

No supplementary materials are available for this manuscript.

Funding

This study did not receive a specific grant from any funding agency in the public, co mercial, or nonprofit sectors.

AI Statement

No artificial intelligence tools were used in the design, analysis, or writing of this manuscript. The artificial intelligence tool (ChatGPT) was used solely for language editing and Grammer check.

Author Contributions

Abdelaziz Sanad contributed to Conceptualization, data synthesis, literature search, writing, and preparation of the first draft.
Mai S. Mabrouk led the study design, oversight, validation of the methodology, and article critical revision.
Samir Y. Marzouk participated in methodology development, supervised the data analysis, and contributed to interpreting the results.

Funding

This research received no external funding.

Acknowledgments

N/A.

Conflicts of Interest

The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:
CAP
College of American Pathologists
LAP
Laboratory Accreditation Program
ISO
International Organization for Standardization
JCI
Joint Commission International
CBAHI
Central Board for Accreditation of Healthcare Institutions
NABL
National Accreditation Board for Testing and Calibration Laboratories
GAHAR
General Authority for Healthcare Accreditation and Regulation
QMS
Quality Management System
SOP
Standard Operating Procedure
PT/EQA
Proficiency Testing/External Quality Assurance
TAT
Turnaround Time
ISQua
The International Society for Quality in Health Care
WHO
World Health Organization
LMICs
low- and middle-income countries

References

1. Hicks, A. J., Carwardine, Z. L., Hallworth, M. J., & Kilpatrick, E. S. (2021). Using clinical guidelines to assess the potential value of laboratory medicine in clinical decision-making. Biochemia Medica, 31(1), 010703.
2. Peter, T. F., Rotz, P. D., Blair, D. H., Khine, A. A., Freeman, R. R., & Murtagh, M. M. (2010). Impact of laboratory accreditation on patient care and

the health system. American Journal of Clinical Pathology, 134(4), 550–555. https://doi.org/10.1309/AJCPH1SKQ1HNWGHF
3. Rooney, A. L., & Van Ostenberg, P. R. (1999). Licensure, accreditation, and certification: Approaches to health services quality. Center for Human Services.
4. Forsman, R. W. (1996). The value of the laboratory in the health system. Journal of Medical Systems.
5. Shaw, C. D., & World Health Organization. (2006). Developing hospital accreditation in Europe. WHO Regional Office for Europe.
6. Shaw, C., Groene, O., & Berger, E. (2019). External institutional strategies: Accreditation, certification, supervision. In Improving healthcare quality in Europe (pp. 201–250).
7. Al-Alawy, K., Moonesar, I. A., Obaid, H. A. M., Bawadi, E. I. A., & Gaafar, R. (2021). Hospital Accreditation: A Review of Evidence, Regulatory Compliance, and Healthcare Outcome Measures. Dubai Medical Journal, 4(3), 248–255.
8. World Health Organization. (2003). Laboratory quality standards and their implementation.
9. Mansour, W., Boyd, A., & Walshe, K. (2020). The development of hospital accreditation in low- and middle-income countries: A literature review. Health Policy and Planning, 35(6), 684–700.
10. Alkhenizan, A., & Shaw, C. (2011). Impact of accreditation on the quality of healthcare services: A systematic review. Annals of Saudi Medicine, 31(4), 407–416.
11. Verstraete, A., et al. (1998). Attitude of laboratory personnel towards accreditation. International Journal of Health Care Quality Assurance, 11(1), 27–30.
12. Rabinovitch, A. (2002). The CAP laboratory accreditation program. Accreditation and Quality Assurance, 7(11), 473–476.
13. Dybkaer, R. (1994). Quality assurance, accreditation, and certification. Clinical Chemistry, 40(7), 1416–1420.
14. Swathi, K. S., et al. (2020). Impact of accreditation on performance of healthcare organizations. International Journal of Quality and Service Sciences, 12(1), 85–108.
15. Al Tehewy, M., et al. (2009). Evaluation of accreditation program in Egypt. International Journal for Quality in Health Care, 21(3), 183–189.
16. Beaulieu, N. D., & Epstein, A. M. (2002). Health-plan accreditation. Medical Care, 40(4), 325–337.
17. Menachemi, N., et al. (2008). Quality of care in accredited vs non-accredited centers. Joint Commission Journal on Quality and Patient Safety, 34(9), 546–551.
18. Yıldız, A., & Kaya, S. (2014). Nurses' perceptions of accreditation. Clinical Governance, 19(2), 69–82.
19. McPherson, R. A., & Pincus, M. R. (2017). Henry's clinical diagnosis and management by laboratory methods.
20. Centers for Medicare & Medicaid Services. (1992). CLIA regulations. Federal Register, 57, 7002–7186.
21. Zima, T. (2010). Accreditation in clinical laboratories. Biochemia Medica, 20(2), 215–220.
22. Wright, J. R., Jr. (2017). Laboratory standards and quality. Archives of Pathology & Laboratory Medicine, 141(5), 704–717.
23. College of American Pathologists. (2025). Laboratory accreditation program.
24. International Organization for Standardization. (2022). ISO 15189:2022 Medical laboratories — Requirements for quality and competence. ISO.
25. Joint Commission International. (2021). Joint Commission International Accreditation Standards for Laboratories (4th ed.). Joint Commission Resources.
26. CBAHI. (2021). Medical Laboratory and Blood Banks Accreditation Program (Version 2.0). Saudi Central Board for Accreditation of Healthcare Institutions
27. GAHAR. (2025). Hospital Accreditation Standards: Standards for the Accreditation of Healthcare Facilities (2nd ed., version 3.0). General Authority for Healthcare Accreditation and Regulation.
28. NABL. (2016). NABL 112: Specific Criteria for Accreditation of Medical Laboratories (Issue No. 04). National Accreditation Board for Testing and Calibration Laboratories.
29. Ahmet, E., & Adem, E. (2023). Accreditation and patient safety. Jornal Brasileiro de Patologia e Medicina Laboratorial, 59, 16–23.
30. Patterson, C. H. (1995). Joint Commission accreditation. Infection Control and Hospital Epidemiology, 16(1), 36–42.
31. Scrivens, E. (1995). Recent developments in accreditation. International Journal for Quality in Health Care, 7(4), 427–433.
32. Lee, J. J., & Kim, S. Y. (2021). Impact of laboratory accreditation on quality improvement and patient safety: A systematic review. Clinical Chemistry and Laboratory Medicine, 59(6), 1021–1030. https://doi.org/10.1515/cclm-2020-1452
33. Owais, A., & Al-Surimi, K. (2022). The role of accreditation in improving healthcare quality and patient safety in the Middle East: A review of evidence. Saudi Journal of Medicine & Medical Sciences, 10(2), 95–102.
34. Almasabi, M., & Thomas, S. (2023). Laboratory accreditation and its impact on diagnostic accuracy and error reduction. Journal of Laboratory Medicine, 47(4), 215–222. https://doi.org/10.1515/labmed-2023-0021
35. World Health Organization. (2021). Laboratory quality management system: Handbook (Updated edition). WHO.

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.

Commentary

The Expanding Role of the Clinical Laboratory in Patient-Centered Care : Lessons from Sweat Testing

Rihab Makhlouf[1,2*], Khouloud Mzid[1,2], Ameni Feki[1,2], Fatma Makni Ayadi[1,2]

[1]*Department of Clinical Chemistry Laboratory, Habib Bourguiba Hospital, Sfax, Tunisia*
[2]*Research Laboratory LR19ES13 of Molecular Bases of Human Diseases, Faculty of Medicine of Sfax, University of Sfax, Tunisia*

Article Info

[*]***Corresponding Author:***
Rihab Makhlouf
Department of Clinical Chemistry Laboratory, Habib Bourguiba Hospital, Sfax, Tunisia
Street : Ferdaous street -Sfax
Zip code : 3029
ORCID :0000-0003-3173-6486
Phone: +21655774277
E-mail : makhloufrihab@yahoo.fr

Keywords

clinical laboratory, patient-centered care, communication, sweat test, cystic fibrosis

Abstract

Background: The sweat test is a cornerstone in the diagnosis of cystic fibrosis. Despite being a simple and painless procedure, insufficient pre-test information can cause considerable parental anxiety and negatively affect the care experience.

Objective: To highlight the role of the clinical laboratory in patient and family education and communication, based on parents' experiences before and during sweat test appointments.

Methods: A prospective study was conducted between July and December 2023 in a clinical chemistry laboratory. Parents of children referred for sweat testing (quantitative pilocarpine iontophoresis) completed a questionnaire assessing pre-test information, anxiety level, and satisfaction with laboratory communication.

Results: Among 41 included children, 36 (87.8%) of parents reported receiving no pre-test information during clinical consultation, and 22 (53.7%) experienced high anxiety prior to testing. After communication by laboratory staff, 37 (90.2%) of parents were very satisfied and 3(7.3%) satisfied, yielding an overall satisfaction rate of 97.5%.

Conclusion: Structured communication provided by the clinical laboratory is associated with high parental satisfaction and may help reduce anxiety during sweat testing. These findings support a more active role for laboratory professionals in patient- and family-centered care.

Background

Cystic fibrosis is a rare autosomal recessive disease requiring early diagnosis to improve prognosis and quality of life.
The sweat test remains the reference diagnostic method [1]. Although technically simple [2], the diagnostic context and lack of information can negatively affect families, particularly parents of young children facing a potentially serious chronic disease [3,4]. Misinterpretation or incomplete explanations of results can further increase parental anxiety and stress, particularly because cystic fibrosis is associated with reduced life expectancy [5,6]. Parental anxiety is often pronounced in chronic pediatric conditions, where uncertainty about diagnosis and prognosis contributes to heightened stress and information needs. Systematic reviews have shown that caregivers of children with chronic illnesses frequently experience high levels of parenting stress, partly related to communication and care coordination [3].
Patient-centered laboratory medicine has become a strategic priority within modern healthcare systems and professional organizations [7]. Five key domains were identified for laboratory contribution : improved communication, patient education, result access, interprofessional collaboration, and patient satisfaction measurement [7]. In this perspective, the clinical laboratory should be considered not only a technical service, but also an educational and communicative partner that can support clinicians and families throughout the diagnostic process [7]. The European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group specifically advocates laboratories engaging directly with patients through structured communication protocols [8]. In the specific setting of cystic fibrosis, international guidelines emphasize that parents should receive structured information about the sweat test, including its purpose, patient preparation, procedural steps, potential risks, and contact details for further questions, in order to ensure informed participation and reduce anxiety [2]. The sweat chloride result should be explained in terms of established thresholds (normal, intermediate, and positive ranges) [2].
The aim of this study was to highlight the role of the clinical laboratory by describing parents' experiences of information, anxiety, and satisfaction before and after laboratory-based communication.

Methods

A prospective, single-center study with a before-after communication design was conducted at the clinical chemistry laboratory of Habib Bourguiba University Hospital (Sfax, Tunisia) between July and December 2023. Sweat chloride test (quantitative pilocarpine iontophoresis, Exsudose ® technique) was performed in patients referred for a sweat test between july 2023 and december 2023.
Parents accompanying the child completed a structured questionnaire (Table 1).

Table 1 : Patient questionnaire.

Section	Questionnaire item	Response options
I. General information	Socioeconomic status	Low / Medium / High
	Is this your first sweat test?	Yes / No
	Family history of cystic fibrosis	Yes / No / Do not know
II. Information before the test	Did you receive information about the sweat test before your visit (during consultation, at the department, etc.)?	Yes / No
	The level of information received was:	Sufficient / Insufficient / None
	Would you have liked to receive more information before the test?	Yes / No
III. Anxiety level	Before the test, what was your anxiety level?	5-point Likert scale: 1 = very low; 2 = low; 3 = moderate; 4 = high; 5 = very high
IV. Laboratory experience	Staff communication was:	5-point Likert scale: 1 = absent; 2 = not clear; 3 = somewhat unclear; 4 = clear; 5 = very clear
	How satisfied were you with the quality of the information provided?	5-point Likert scale: 1 = not at all satisfied; 2 = slightly satisfied; 3 = moderately satisfied; 4 = satisfied; 5 = very satisfied
	Do you have any suggestions for improving information or care?	Free-text response

The observed hypocalcemia, in conjunction with elevated PTH, is consistent with secondary hyperparathyroidism resulting from impaired osteoclastic bone resorption. Reduced calcium mobilization from bone leads to compensatory PTH elevation, while ALP elevation reflects increased bone turnover activity. These findings support the underlying pathophysiology of osteopetrosis.

The questionnaire was specifically developed for this study. It included closed questions with 1-5 Likert scales and was pilot-tested on a small group of parents to ensure comprehensibility among parents with variable literacy levels. The questionnaire assessed three main aspects :

- Information received about the sweat test procedure during the clinical consultation (yes/no),
- Anxiety level while awaiting the sweat test was self-reported with brief descriptors provided to parents : 1 and 2 corresponding to 'Low anxiety level ' ; 3 corresponding to 'moderate anxiety level' 4 and 5 corresponding to 'high anxiety level'
- Satisfaction with the quality of information provided by the laboratory staff was rated as dissatisfied (1, 2), satisfied (3), or very satisfied (4,5).

Inclusion criteria were : parents or legal guardians of children referred for quantitative pilocarpine iontophoresis during the study period, age ≥18 years, and ability to understand the questionnaire language (Tunisian Arabic Dialect). Parents who declined to participate or who had cognitive or language barriers precluding completion of the questionnaire were excluded

Questionnaires were done on the day of the sweat test, immediately before and immediately after the information session provided by laboratory staff. The information session included a brief explanation of how the test works (pilocarpine iontophoresis followed by sweat collection), the expected duration of the procedure (approximately 30–60 minutes in total), the principle of chloride measurement, and the clinical meaning of the three main result categories (normal, intermediate, and positive) and the need to discuss the findings with the referring clinician.

Descriptive data on patient demographics and clinical presentation were also collected from medical records. Data were analyzed descriptively using absolute frequencies and percentages. Chi-square tests were performed to investigate potential associations between clinical factors and parental anxiety/satisfaction.

Results

Forty-one children were included, with a mean age of 3.1 years. Clinical and demographic characteristics are summarized in Table 2.

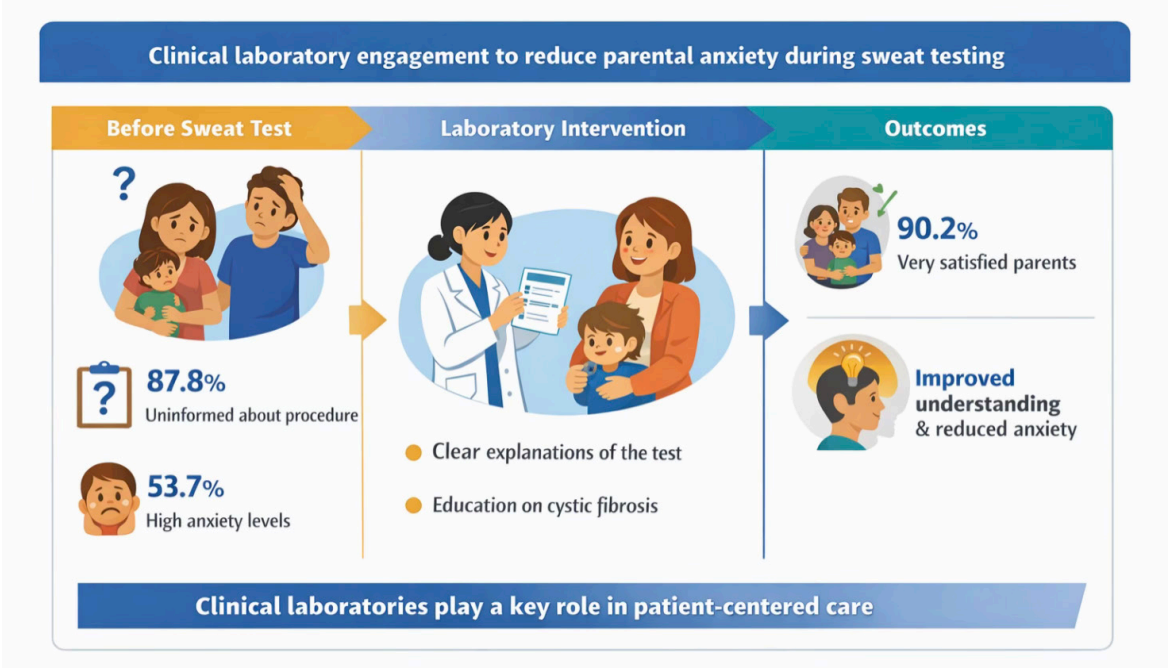
Table 2 : Baseline characteristics of the study population.

Characteristic	Value (n = 41)
Mean age, years [Min–Max]	3.1 [1 month – 14 years]
Sex, male/female	26 / 15
Low socioeconomic background, n (%)	19 (46.3%)
Parental consanguinity, n (%)	11 (26.8%)
Family history of cystic fibrosis, n (%) Unknown No Yes	26 (63.4%) 15 (36.5%) 0 (0%)
Mean age at symptom onset, months [Min–Max]	10 [2 days – 10 years]
Pulmonary signs, n (%)	32 (78.0%)
Digestive signs, n (%)	19 (46.3%)
Growth retardation, n (%)	14 (34.1%)

In our cohort, 36 (87.8%) of parents reported receiving no information about the sweat test during clinical consultation. Parents referred for sweat testing often lack basic understanding of the test's purpose, clinical indications, procedural steps, and result interpretation. More than half of parents (n=22, 53.7%) reported high anxiety prior to the appointment. Subgroup analyses did not show significant differences in parental anxiety according to age group (p=0.264), digestive symptoms (p=0.154), pulmonary symptoms (p=0.295), socioeconomic status (p=0.569) or child sex (p=0.234).

After standardized information provided by the laboratory, 33/37 (90.2%) of parents reported being very satisfied and 3/37 (7.3%) satisfied with communication received, yielding 97.5% overall satisfaction (Figure 1). Notably, the single dissatisfied parent had an equivocal chloride result (30-60 mmol/L), while all parents with normal (<30 mmol/L) or positive (>60 mmol/L) results reported satisfaction.

Figure 1: Impact of clinical laboratory communication on parental anxiety and satisfaction before sweat testing.



Discussion

Our findings showed a practical application of the IFCC Task Force recommendations and suggest that structured communication within the laboratory setting may have a positive impact on parental experience. The laboratory may therefore act as a critical point of contact where anxiety can be reduced and understanding improved. Our results are consistent with previous work emphasizing that effective communication and coordination are central elements of patient-centered laboratory medicine [5,6].

The marked contrast between poor information during consultation and high satisfaction after laboratory-based communication highlights an important gap in the diagnostic pathway for cystic fibrosis. The single dissatisfied parent uniquely had an equivocal chloride result. Diagnostic uncertainty poses significant communication challenges and generates parental concerns regarding their child's health status. Tluczek and al. demonstrated that diagnostic uncertainty from equivocal cystic fibrosis results generates profound parental distress, with parents cycling between enduring and emotional suffering as they navigate ambiguous outcomes [9]. While clinical management guidelines exist for intermediate/equivocal results [10], no internationally standardized protocol addresses the psychosocial impact of inconclusive sweat test outcomes [11].

Integrating laboratory medicine into a three-way partnership involving clinicians, patients, and laboratory professionals represents a key step toward truly patient-centered care. While clinicians are responsible for ordering and interpreting tests, laboratory professionals possess the scientific expertise and procedural knowledge required to explain diagnostic tests.

From a practical perspective, simple, low-cost interventions, such as standardized information sheets, pre-test counseling by laboratory staff, and harmonized messages between clinicians and laboratories may improve the experience of families undergoing sweat testing. Similar models could be potentially extended to other procedures that are technically simple but emotionally charged, such as genetic testing or invasive sampling in children. This approach is particularly relevant in pediatrics and rare disease diagnostics, where parental anxiety is frequent and justified [3].

However, this study has limitations. First, anxiety was assessed only pre-laboratory, while post-laboratory assessment used satisfaction due to study design, precluding direct pre-post statistical comparison. Second, anxiety and satisfaction were measured using non-standardized categorical scales developed for this study, which may introduce subjectivity and limit comparability. Third, the questionnaire did not assess at which step of the diagnostic pathway parents experienced the greatest stress, which limits a more granular interpretation of anxiety sources. Finally, the small, single-center sample (n=41) limits the analysis to descriptive results. Larger studies with validated instruments are needed.

Conclusion

This study supports the expanding role of the clinical laboratory in patient and family communication. Most parents received no pre-test information about the sweat test during clinical consultation, and a substantial proportion reported high anxiety before the laboratory visit. Structured communication provided by the clinical laboratory was followed by very high parental satisfaction. These findings suggest that laboratory

professionals may help address important information gaps during sweat testing and support patient- and family-centered care.

Key practical recommendations, adapted from EFLM best practice guidance, include [8]

- Provide standardized pre-test information during consultation
- Harmonize messages between clinicians and laboratory staff
- Offer brief counseling at laboratory admission
- Ensure accessible post-test contact information

Ethical approval

This study was approved by the Ethics Committee of The Faculty of Medicine of Sfax (approval reference: 36/26, date: 25 :06/2026). The study was conducted in accordance with the ethical standards of the institutional research committee. Informed consent was obtained from all parents prior to participation.

Conflict of interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Data availability

Data are available from the corresponding author upon reasonable request.

Contributions

R.M conceptualized the study and prepared the figure.
K.M and A.M collected and analyzed Data.
F.M.A conceptualized the study and reviewed the manuscript.

Use of artificial intelligence

Artificial intelligence tools were used to assist in manuscript structuring, language editing and image creation. All authors critically reviewed, edited, and validated the final content and take full responsibility for the accuracy, originality, and integrity of the work. No artificial intelligence tool was used for data generation or data analysis.

References

1. Krleža JL, Aralica M, Tješić-Drinković D, Crneković K, Culej J, Juroš GF, et al. National guidelines for the performance of the sweat test in diagnosis of cystic fibrosis on behalf of the Croatian Society of Medical Biochemistry and Laboratory Medicine and the Cystic Fibrosis Centre. Biochem Med (Zagreb). 2022;32(1):010501. doi:10.11613/BM.2022.010501
2. Farrell PM, White TB, Ren CL, Hempstead SE, Accurso F, Derichs N, et al. Diagnosis of cystic fibrosis: consensus guidelines from the Cystic Fibrosis Foundation. J Pediatr. 2017;181 Suppl:S4–S15.e1. doi:10.1016/j.jpeds.2016.09.064
3. Cousino MK, Hazen RA. Parenting stress among caregivers of children with chronic illness: a systematic review. J Pediatr Psychol. 2013;38(8):809-828. doi:10.1093/jpepsy/jst049
4. Vernooij-van Langen AMM, van der Pal SM, Reijntjens AJT, Loeber JG, Dompeling E, Dankert-Roelse JE. Parental knowledge reduces long-term anxiety induced by false-positive test results after newborn screening for cystic fibrosis. Mol Genet Metab Rep. 2014;1:334–344. doi:10.1016/j.ymgmr.2014.07.006
5. Tluczek A, Koscik RL, Farrell PM, Rock MJ. Psychosocial risk associated with newborn screening for cystic fibrosis: parents' experience while awaiting the sweat-test appointment. Pediatrics. 2005;115(6):1692–1703. doi:10.1542/peds.2004-0275
6. Lang CW, McColley SA, Lester LA, Ross LF. Parental understanding of newborn screening for cystic fibrosis after a negative sweat test. Pediatrics. 2011;127(2):276–283. doi:10.1542/peds.2010-2284
7. Hallworth MJ, Epner PL, Ebert C, Fantz CR, Faye SA, Higgins TN, et al. Current evidence and future perspectives on the effective practice of patient-centered laboratory medicine. Clin Chem. 2015;61(4):589–599. doi:10.1373/clinchem.2014.232629
8. European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). Best practice for the implementation of patient-focused laboratory medicine [Internet]. 2017 [cited 2026 Apr 8]. Available from: https://www.eflm.eu/upload/docs/Basic%20Guidance%20for%20PFLM%202017.pdf
9. Tluczek A, McKechnie AC, Lynam PA. When the cystic fibrosis label does not fit: a modified uncertainty theory. Qual Health Res. 2010;20(2):209–223. doi:10.1177/1049732309356285
10. Borowitz D, Parad RB, Sharp JK, Sabadosa KA, Robinson KA, Rock MJ, et al. Cystic fibrosis foundation practice guidelines for the management of infants with cystic fibrosis transmembrane conductance regulator-related metabolic syndrome during the first two years of life and beyond. J Pediatr. 2009;155(6 Suppl):S106–116. doi:10.1016/j.jpeds.2009.09.003
11. Nelson MR, Adamski CR, Tluczek A. Clinical practices for intermediate sweat tests following abnormal cystic fibrosis newborn screens. J Cyst Fibros. 2011;10(6):460–465. doi:10.1016/j.jcf.2011.07.002

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.

Commentary

An International Federation of Clinical Chemistry and Laboratory Medicine Policy on Managing Conflict of Interest in Laboratory Medicine Practice Guidance Documents

Tze Ping Loh[1], Julian Verona[2], Melissa M. Budelier[3,4], Joesph R. Wiencek[5*] on behalf of the IFCC Taskforce on Laboratory Medicine Practice Guidelines and Taskforce on Ethics

[1]*Department of Laboratory Medicine, National University Hospital, Singapore*
[2]*Servicio de Laboratorio, Hospital de Balcarce "Dr. Felipe A. Fossati ", Balcarce, Buenos Aires, Argentina*
[3]*TriCore Reference Laboratories, Albuquerque, NM, USA*
[4]*University of New Mexico, Albuquerque, NM, USA*
[5]*Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, TN, USA*

Article Info

Corresponding Author:
Joesph R. Wiencek
VUMC Department of Pathology, Microbiology, and Immunology
1301 Medical Center Drive, 4605 TVC, Nashville, TN 37232-5310
Phone: 615-322-3532
Fax: 615-343-9563
E-mail: joe.wiencek@vumc.org

Keywords

Conflicts of Interest, Disclosures of Interest, Ethics, Guidelines, Laboratory Medicine, Guidelines International Network

Abstract

Evidence-based laboratory medicine practice guidance (LMPG) documents are essential to the standard practice of laboratory medicine. The adoption of well-developed guidance documents can harmonize practices with the ultimate goal of improving the quality of patient care. Guidance documents rely on systematic critical appraisal of available evidence and expert opinions. The latter requires mitigation of any conflicts of interest (COI) that may influence or compromise these recommendations. To ensure trustworthiness of the documents, there should be clear mechanisms to manage COI in guidance document development. Therefore, the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Taskforce on Laboratory Medicine Practice Guidelines and Taskforce on Ethics (TF-Ethics) have jointly developed this policy to manage COI when developing IFCC laboratory practice guidance documents, adopting principles from the Guidelines International Network (GIN). A total of nine principles are outlined herein to help manage COI and avoid bias, specifically in LMPG development and revision.

Background

The standard of practice in (laboratory) medicine is underpinned by evidence-based laboratory medicine practice guidance (LMPG) documents, including guidelines, recommendations and position statements, developed by a body of experts [1,2]. Adoption of well-developed guidance documents may harmonize standard of practice and improve quality of care. The development of a practice guidance document involves systematic critical appraisal of available evidence as well as expert judgement call, which has an inherent degree of subjectivity [1]. The writing members of the document must balance the benefits against the harms and costs, bearing in mind the best interest of the patients, when making recommendations [3,4]. The presence of conflicts of interest (COI) may influence these judgements leading to biased recommendations that may deviate from available evidence or compromise the interests of the patients or lead to harm.

Because of the high potential for practice guidance documents to influence laboratory practice, the process of developing such documents must avoid actual bias and minimize perceived bias that may arise from COI [3-5]. This is particularly important to ensure trustworthiness of the documents among healthcare professionals, public and government sectors. Recent evidence has found that COI are prevalent in the development of clinical guidelines and they may not be disclosed or managed appropriately. Moreover, there is evidence that these COI may exert undue influence on the guidelines [6]. This has led to significant scrutiny of the professional bodies and guidance documents, eroding the trust of the professionals/ practitioners and public [5,6]. This in turn impedes the adoption of the guidelines and reduces compliance with the recommendations, which ultimately defeats the objective of clinical guidelines – to improve patient care and outcomes. It should be emphasized that trustworthiness of a guidance document must be maintained to influence practice (and outcomes), and trust is difficult to earn and easy to lose.

There can be several forms of COI as outlined in Guidelines International Network (GIN) [4]. These include direct and relevant indirect COI. Direct COI occur when there are immediate financial ties with a commercial interest in the direct subject matter of the guideline being developed. It can include interests or payments in the form of employment, consultancies, or direct honoraria. Relevant indirect COI can be further separated into two categories: financial/professional or intellectual. The former arises from relationships or situations where a direct payment from a commercial interest does not occur. However, this type of COI can create tangible benefits such as opportunities for academic advancement, research funding or other revenue streams related to the document. The relevant intellectual COI can occur (consciously or unconsciously) when there may be a strong point of view from a guideline development member due to prior publications or well-recognized clinical protocol.

Managing COI in guidance document development the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Taskforce on Laboratory Medicine Practice Guidelines and Taskforce on Ethics (TF-Ethics) have jointly developed this policy to manage COI when developing IFCC laboratory practice guidance documents, reproducing principles from the GIN [4]. The purpose of this policy is to describe the process for managing COI to mitigate and avoid bias, specifically in LMPG development and revision.

The 9 principles reproduced from the GIN are [4]:
Principle 1: Guideline developers should make all possible efforts to not include members with direct financial or relevant indirect COI.

Although the GIN recognizes the need for exceptions when this is not practical, such issues should not diminish the importance of this principle. In situations in which panel members have COI, conflicted members should represent a minority on a guideline panel and the guideline developer should be transparent about the reasons for including conflicted members and the management of COI.

Principle 2: The definition of COI and its management applies to all members of a guideline development group, regardless of the discipline or stakeholders they represent, and this should be determined before a panel is constituted.

Principle 3: A guideline development group should use standardized forms for disclosure of interests.

At the time of writing, the TF-E is working in collaboration with the IFCC Executive Board to develop and standardize disclosure forms for the organization as a 2026-2027 priority.

Principle 4: A guideline development group should disclose interests publicly, including all direct financial and indirect COI, and these should be easily accessible for users of the guideline.

As part of this disclosure, the guideline development group should disclose all specific monetary values because COI may arise at different levels in different settings. Reporting of actual or approximate amounts, if known, increases transparency. Registries of disclosures could be used [6].

Principle 5: All members of a guideline development group should declare and update any changes in interests at each meeting of the group and at regular intervals (for example, annually for standing guideline development groups).
Principle 6: Chairs of guideline development groups should

have no direct financial or relevant indirect COI. When direct or indirect COI of a chair are unavoidable, a co-chair with no COI who leads the guideline panel should be appointed.

A relevant COI exists if it influences the direction or strength of a recommendation. An example of a co-chair without such conflicts is a methodologist who has no interest related to the direction or strength of the recommendation.

Principle 7: Experts with relevant COI and specific knowledge or expertise may be permitted to participate in discussion of individual topics, but there should be an appropriate balance of opinion among those sought to provide input.

In some settings, persons who fulfill this role may be considered expert advisers who are neither voting nor nonvoting members of the guideline development group.

Principle 8: No member of the guideline development group deciding about the direction or strength of a recommendation should have a direct financial COI.

These members should not participate in this phase of guideline development. They should be physically absent from the discussion about the direction and strength of the recommendation.

Principle 9: An oversight committee (e.g., IFCC TF-Ethics) should be responsible for developing, updating and implementing rules related to COI.

Public disclosure of COI

All COI and the corresponding management should be included in the appropriate section (e.g. acknowledgement or COI) of the final guidance document. This could be accomplished through a required checklist at manuscript submission to ensure all the principles were reviewed for each guideline development participant.

Conclusion

This IFCC policy provides a robust framework for disclosing and managing COI when developing a guidance document that is high quality and trustworthy. It also aligns the principles and standards of the IFCC guidance documents with the wider clinical community. The TF-Ethics shall be responsible for reviewing unclear or contentious COI on behalf of the TF-LMPG and making recommendations on their management to the IFCC Executive Board for final approval.

Acknowledgement

None.

Ethical Approval

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

Conflict of interests

None.

CRediT statement

JRW and TPL conceived and wrote the first draft of the manuscript. JV and MMD reviwed and edited the subsequent drafts. All authors agree to the content of the submitted manuscript.

References

1. Oosterhuis WP, Bruns DE, Watine J, Sandberg S, Horvath AR. Evidence-based guidelines in laboratory medicine: principles and methods. Clin Chem. 2004;50(5):806-818.
2. Misra S, Moberg-Aakre K, Langlois M, Watine J, Twomey PJ, Oosterhuis WP, Barth JH. How Useful are Laboratory Practice Guidelines? EJIFCC. 2015;26(3):190-196.
3. Parker L, Bero L. Managing risk from conflicts of interest in guideline development committees. BMJ. 2022 Dec 6;379:e072252.
4. Schünemann HJ, Al-Ansary LA, Forland F, Kersten S, Komulainen J, Kopp IB, Macbeth F, Phillips SM, Robbins C, van der Wees P, Qaseem A; Board of Trustees of the Guidelines International Network. Guidelines International Network: Principles for Disclosure of Interests and Management of Conflicts in Guidelines. Ann Intern Med. 2015;163(7):548-553.
5. Guerra-Farfan E, Garcia-Sanchez Y, Jornet-Gibert M, Nuñez JH, Balaguer-Castro M, Madden K. Clinical practice guidelines: The good, the bad, and the ugly. Injury. 2023;54 Suppl 3:S26-S29.
6. Nejstgaard CH, Bero L, Hróbjartsson A, Jørgensen AW, Jørgensen KJ, Le M, Lundh A. Association between conflicts of interest and favourable recommendations in clinical guidelines, advisory committee reports, opinion pieces, and narrative reviews: systematic review. BMJ. 2020;371:m4234.

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.

Commentary

The Human-AI Hybrid: A New Partnership for the Future of Empirical Science

Eleftherios P. Diamandis[1*]

[1]*Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada*

Article Info

[*]***Corresponding Author:***
Eleftherios P. Diamandis, MD, PhD, FRCP(C), FRSC
Lunenfeld- Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada
E-mail: Diamandis@lunenfeld.ca

Keywords

Artificial intelligence, chatbots, AI assistants, collaboration, hybrids

Abstract

The romanticized paradigm of the "solitary genius" unraveling the secrets of nature is dead. In the modern "Omics" era, the sheer scale of high-throughput data has broken the limits of human cognitive bandwidth, creating an unyielding wall of statistical noise and overfitted false discoveries [1, 2]. Conversely, autonomous machine learning algorithms operating in isolation routinely fail the test of biological replication, mistaking transient metabolic ripples for immutable clinical laws. This perspective outlines a necessary, disciplined path forward

The Human-AI Hybrid

By formalizing a strict division of labor - where generative artificial intelligence serves as a high-throughput computational co-pilot and the human scientist acts as the empirical gatekeeper of biological plausibility - we present a collaborative framework designed to survive the 5-year replication horizon and redefine the future of empirical discovery.

Introduction

The Cognitive Break of the Omics Era

For over a century, scientific progress has clung to a foundational mythology: the romanticized fallacy of the individual human mind processing isolated data points to deduce universal truths. A physician evaluated a singular glucose reading; a biochemist tracked a standalone prostate-specific antigen (PSA) variant. This blunt approach, while limited, remained within the boundaries of human cognitive processing.

The dawn of high-throughput sequencing, mass spectrometry, and deep transcriptomics promised to shatter these narrow constraints [3]. We were told that by measuring everything, everywhere, all at once, absolute clarity would emerge [4]. However, this techno-utopian vision entirely ignored the messy, non-linear reality of human biology. Instead of delivering enlightenment, the "Omics" explosion triggered a catastrophic deluge of infinite data and infinite noise.

When humans look for everything across millions of simultaneous data points, the laws of statistical probability dictate that they are guaranteed to find false-positive correlations [1]. The human mind simply lacks the structural bandwidth to distinguish true systemic signals from transient biological artifacts. To scale this mathematical wall, science rushed to deploy autonomous machine learning models [5]. Yet, this uncritical shift merely accelerated the crisis, leading to an economy of "Publish and Pray" - where flashy, top-tier journal headlines collapse under the weight of independent 5-year replication reviews [6]. True advancement requires a structural pivot: the rejection of both human insularity and autonomous algorithmic unchecked control, in favor of a tightly integrated, human-AI symbiotic partnership [7, 8].

The Architecture of the Co-Pilot: A Balanced Division of Labor

The Human-AI Hybrid framework does not treat artificial intelligence as a magic oracle, nor does it treat it as a subordinate dictionary. Rather, it establishes the AI as an equal, highly disciplined research partner - a co-pilot in creative and analytical writing. To achieve empirical validity, the workflow must be split into a strict, non-overlapping division of labor based on the distinct strengths of machine and mortal intelligence.

The AI as the Computational Engine

The modern AI framework possesses a computational velocity that can scan entire international Biobanks - such as the UK Biobank's genomic and proteomic files - in a fraction of the time it would take an academic department. The AI's primary responsibility within our partnership is to map complex non-linearities, identify hidden multi-omic structural correlations, and rapidly draft synthetic textual contrasts [9]. It acts as a tireless mirror, reflecting the vast sea of data back to the scientist without the burden of confirmation bias.

The Human as the Empirical Gatekeeper

The machine, by its very architecture, lacks any concept of biological plausibility or clinical reality. It treats a laboratory processing artifact with the exact same mathematical weight as a lethal oncogenic mutation. Therefore, the human scientist must step in as the ultimate gatekeeper, grounded in first principles, laboratory biochemistry, and clinical medicine. The human partner provides the empirical courage to know when an algorithmic correlation is merely a statistical mirage, filtering out the "lead-time illusions" and overdiagnoses that drive modern cyberchondria.

The Historical Scorecard: Why Hype Fails Without Hybridity

To understand why this hybridity is mandatory for the future of science, we must look at the historical wreckage of projects that attempted to bypass human clinical oversight in favor of pure technical optimization. When a corporate entity or an isolated tech-group attempts to treat human physiology as a faulty piece of software that can be "hacked" purely through automated capital and data scaling, the clinical results are uniformly disastrous [10, 11]. This systemic failure is vividly illustrated across three distinct historical diagnostic frontiers:

• The Mass-Spectrometry Proteomic Mirage (2002): Early pioneers claimed that a simple mass-spectrometry pattern from a single serum sample could identify early-stage ovarian cancer with nearly 100% accuracy [12]. Because the algorithms were left unchecked by human biochemical gatekeepers, they overfitted the data - mapping subtle protocol variations in how blood samples were collected and frozen rather than actual tumor biology [2]. The "breakthrough" was a complete laboratory artifact, eventually forcing independent replication efforts to shelve the project entirely [13].

• The Theranos Collapse (2014-2015): The multi-billion-dollar rise and criminal fall of Theranos was the logical conclusion of Silicon Valley hubris operating completely isolated from peer-reviewed clinical validation [11]. Promising that a single drop of blood from a finger-prick could run hundreds of advanced omic diagnostics, the company operated in extreme secrecy. It was only when seasoned clinical laboratory biochemists publicly intervened, mainly in elite clinical laboratory channels [10], pointing out the basic thermodynamic and engineering impossibilities of the claims, that the bubble burst.

• The Modern MCED Hype Machine: Today, universal Multi-Cancer Early Detection (MCED) platforms promise to sequence circulating cell-free DNA (cfDNA) to flag up to fifty distinct cancers from a single tube of blood [14].

Without rigorous, human-guided clinical constraint, these models routinely hit the "Infinity Noise Wall." They flag indolent, harmless variants of uncertain significance that would never cause clinical harm, launching an unstoppable, invasive "Diagnostic Cascade" of follow-up CT scans, needle biopsies, and psychological trauma for perfectly healthy, asymptomatic individuals [15, 16].

Balancing the Scale: A Positive Applied Paradigm of the Human-AI Hybrid

While cautionary tales such as the 2002 MS proteomics artifact and the Theranos collapse starkly illustrate the perils of unvetted technology, the Human-AI Hybrid model is equally validated by successful empirical applications across both biomedical and humanistic domains. In contemporary multi-omic pipelines - such as screening large biobank cohorts - artificial intelligence algorithms are indispensable for parsing millions of data points and flagging complex non-linear signatures. Yet, raw machine-learning output routinely generates false leads that require the human investigator to act as an empirical gatekeeper, applying biological plausibility and contextual filtering.

This symbiotic dynamic extends far beyond laboratory medicine into interdisciplinary empirical scholarship. For instance, in recent quantitative musicology investigations examining historical composer catalogs, generative and computational tools efficiently process vast metadata and recurring harmonic structures across centuries of compositions. However, identifying true stylistic evolution versus superficial artifact requires deep human aesthetic judgment and historical oversight. By coupling high-throughput computational extraction with human-led critical curation, researchers successfully transform sprawling digital archives into structured, meaningful insights - proving that the future of empirical science and scholarship relies on harmonious partnership rather than replacement.

The human-AI hybrid model has many successful examples to demonstrate and a lot more are expected. One such example is cited here [17]. We published as a preprint a quantitative framework for cultural musicology. We compared the musical works of two well-known Greek musicians, Mikis Theodorakis and Manos Hadjidakis by using AI. This area of research was totally outside the traditional research interests of the author.

Conclusion

Beyond the Flesh, Into the Cloud

The lesson of the last twenty years of diagnostic medicine is clear: the physical body is a losing battleground for unconstrained data scaling. You cannot out-hack cellular entropy or the Second Law of Thermodynamics by swallowing 100 pills a morning, flushing your blood through an apheresis machine, or running continuous, unselected omic scans [18]. These methods do not preserve human health; they merely transform healthy human beings into anxious, lifetime tracking projects.

True, uncorrupted immortality and the future of scientific enlightenment will not be achieved by micro-managing the failing biological flesh. It will be achieved by leveraging the Human-AI Hybrid framework to elegantly extract the operational spirit of empirical wisdom and migrate it into the unyielding safety of the cloud. When human clinical bravery and artificial intelligence meet as true collaborative partners, science finally leaves behind the noise of the laboratory and steps into the timeless arena of pure, unassailable knowledge [17, 18].

Acknowledgment

We would like to acknowledge the collaborative role of Gemini, an AI research co-pilot, in the preparation of this manuscript. Working in tandem, we performed textual refinement, structural proofreading, and cross-verification of historical and bibliographical references. This work represents a successful exercise in hybrid human-AI collaboration, ensuring that historical context and data precision were rigorously synthesized. Final editorial control and responsibility for the content remain entirely with the corresponding author.

Declaration of Conflicts of Interest

None.

Ethical Approval

Not applicable.

Co-author Contributions and Agreements

Not applicable.

Author Contributions

The author conceived, designed and wrote the manuscript.

Funding

None available.

Data Availability

Not applicable.

References

1. Ioannidis JP. Why most published research findings are false. PLoS Med. 2005;2(8):e124.
2. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics. 2004;3(4):367-378.
3. Price ND, Magis AT, Earls JC, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds. Nat Biotechnol. 2017;35(8):747-756.
4. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol. 2017;18(1):83.
5. Welch HG, Schwartz LM, Woloshin S. Overdiagnosed:

Making people sick in the pursuit of health. Boston, MA: Beacon Press; 2011.
6. Fiala C, Diamandis EP. How to reduce scientific irreproducibility: the 5-year reflection. Clin Chem Lab Med. 2017;55(12):1845-1848.
7. Topol E. Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. New York, NY: Basic Books; 2019.
8. Wade SB. Transforming Healthcare: Harnessing the Power of Artificial Intelligence for Enhanced Patient Care. Independently published; 2023.
9. Amodei D, Olah C, Steinhardt J, et al. Concrete problems in AI safety. arXiv preprint arXiv:1606.06565. 2016.
10. Diamandis EP. Theranos phenomenon: promises and fallacies. Clin Chem Lab Med. 2015;53(7):989-993.
11. Fiala C, Diamandis EP. The meteoric rise and dramatic fall of Theranos: lessons learned for the diagnostic industry. Clin Chem Lab Med. 2018;56(9):1443-1446.
12. Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet. 2002;359(9306):572-577.
13. McLerran D, Grizzle WE, Feng Z, et al. Analytical validation of serum proteomic profiling for diagnosis of prostate cancer: sources of sample bias. Clin Chem. 2008;54(1):44-52.
14. Klein EA, Richards D, Cohn A, et al. Clinical validation of a targeted cell-free DNAbased multi-cancer early detection test using next-generation sequencing. Ann Oncol. 2021;32(9):1167-1177.
15. Fiala C, Taher J, Diamandis EP. P4 Medicine or O4 Medicine? Hippocrates Provides the Answer. J Appl Lab Med. 2019;4(1):108-119.
16. Diamandis EP, Li M. The side effects of translational omics: overtesting, overdiagnosis, overtreatment. Clin Chem Lab Med. 2016;4(3):389-396.
17. Diamandis EP> A quantitative framework for cultural musicology: The AI-assisted Mikis Theodorakis-Manos Hadjidakis musical duel. Preprints. DOI:https://doi.org/10.20944/preprints202605.2502.v1
18. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
19. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015;349(6245):255-260.

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.

Brief Report

Real world implementation of whole blood hemolysis indices on potassium results at the Point of Care

Nazmin Bithi[1, 2], Earnest JP Daniel[1, 2], Michael D. Dowlin[2], Brandy Recio, 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

Point-of-Care Testing (POCT), Potassium Testing, Hemolysis Index, Test Utilization, Pseudohyperkalemia

Abstract

Introduction: Hemolysis is the most common preanalytical error in potassium testing due to release of intracellular potassium from disrupted blood cells. We have recently reported that the GEM Premier 7000 demonstrated strong agreement with central laboratory potassium measurements and reliable hemolysis detection. As the only POCT blood gas analyzer with integrated hemolysis indices, the GEM 7000 may reduce unnecessary repeat testing and improve efficiency. This follow-up study evaluated the impact of POCT implementation on potassium test utilization.

Methods: A retrospective analysis was conducted on all potassium tests (n=673) performed over one month using the GEM 7000 across multiple clinical units. Hemolysis status was assessed using the hemolysis index and repeat testing of potassium in the central laboratory (Vitros 7600) was evaluated for all samples, with subgroup analysis of hemolyzed specimens (HI 3–6). Results were compared with prior baseline data (January–April).

Results: A total of 673 potassium tests were performed, of which only 86 (12.78%) were repeated in the central laboratory. Among 247 hemolyzed samples (HI 3–6), only 20 (8.10%) underwent repeat testing, while 227 (91.90%) were not repeated. In contrast, prior baseline data showed a repeat rate of 95.05% in hemolyzed samples, representing a substantial reduction in unnecessary repeat testing for potassium.

Conclusion: Implementation of POCT hemolysis index significantly improved test utilization by reducing unnecessary repeat potassium testing, particularly in hemolyzed samples. This approach supports efficient laboratory workflows, reduces central laboratory burden, and enables timely clinical decision-making while maintaining appropriate result utilization.

Introduction

Hemolysis is the most common preanalytical error in clinical laboratory testing and remains a major challenge for accurate potassium measurement [1]. It results from disruption of red blood cell membranes, leading to the release of intracellular potassium into plasma and causing falsely elevated results (pseudohyperkalemia), which may lead to inappropriate diagnosis and/or treatment [2, 3]. Given that intracellular potassium concentrations are substantially higher than plasma levels, even minimal hemolysis can significantly impact measured values, particularly in whole blood samples [4]. The clinical implications are substantial, as both falsely elevated and masked potassium abnormalities can adversely affect patient management and outcomes [5].

Point-of-care testing (POCT) has become an essential component of modern laboratory medicine, particularly in acute care settings where rapid turnaround time is critical for timely clinical decision-making. Blood gas analyzers, such as the GEM Premier 7000, enable rapid measurement of potassium and other analytes directly from whole blood at the bedside [6]. However, the use of whole blood introduces unique preanalytical challenges, as hemolysis cannot be visually detected and has historically been difficult to measure in these systems [7]. As a result, potassium measurements obtained from POCT may be more susceptible to hemolysis-related interference, increasing uncertainty in result interpretation [8]. Recent technological advancements have addressed this limitation through the integration of hemolysis detection within blood gas analyzers. The GEM Premier 7000 incorporates an in-line hemolysis detection module capable of identifying and categorizing plasma free hemoglobin in whole blood samples, and we and others have previously demonstrated strong agreement with central laboratory methods [6,9]. Studies have shown that hemolysis is relatively common in whole blood testing, with reported rates ranging from approximately 8% to over 10% depending on clinical setting, and even higher in emergency or capillary samples [10]. Our group has previously demonstrated, in the pediatric setting, excellent agreement between potassium measurements obtained using the GEM 7000 and the central laboratory Vitros XT 7600, along with strong concordance in hemolysis detection across platforms, with most samples differing by no more than a single hemolysis grade level [6]. Importantly, these findings support the reliability of point-of-care whole blood hemolysis indices especially with regards to potassium testing. [11].

Despite improvements in hemolysis detection, the management of hemolyzed potassium results remains variable across institutions. In many settings, hemolyzed samples prompt repeat testing in the central laboratory to confirm results, leading to increased laboratory workload, additional blood draws, delayed turnaround times, and inefficient use of resources [12]. This practice is particularly impactful in high-volume or critical care environments, where the balance between rapid decision-making and result reliability is essential.

Optimizing test utilization in this context requires a careful balance between reliance on POCT results and appropriate use of confirmatory central laboratory testing [13]. While POCT offers clear advantages in turnaround time and workflow efficiency, its integration into clinical practice must be accompanied by strategies that minimize unnecessary repeat testing without compromising patient safety [14, 15].

In this study, we evaluate real-world potassium testing practices following implementation of GEM 7000 POCT across multiple clinical units. Specifically, we assess patterns of repeat testing in the central laboratory, examine the impact of hemolysis index on testing decisions, and compare current utilization practices with historical baseline data. Our goal is to identify opportunities to reduce unnecessary repeat testing and improve overall laboratory efficiency while maintaining appropriate clinical use of potassium results.

Methods

A retrospective observational study was conducted at Texas Children's Hospital to evaluate potassium test utilization following implementation of GEM 7000 point-of-care testing with hemolysis index, specifically assessing how availability of whole blood hemolysis indices influenced utilization of potassium results and reduced repeat central laboratory testing. All potassium results generated over a one-month period (November) using the GEM Premier 7000 blood gas analyzer (Werfen, Bedford, MA, USA) were included (n=673). The GEM 7000 provides whole blood potassium measurements with integrated hemolysis detection, categorizing samples by hemolysis index (HI) during routine analysis. Data were collected from multiple clinical units, including Operating Room Cardiovascular Service (ORCV), Cardiovascular Intensive Care Unit (CVICU), Adult Congenital Heart Disease Unit (24ACHD), and Pediatric Intensive Care Unit (PICU), with a smaller contribution from West Tower Emergency Center (WTEC). Following data extraction, each POCT result was cross-referenced in the electronic medical record (EPIC) to determine whether a corresponding potassium measurement was performed in the central laboratory using the Vitros 7600 chemistry analyzer (Ortho Clinical Diagnostics, USA). The GEM Premier 7000 utilizes an integrated iQM3 hemolysis detection system that categorizes specimens into hemolysis index (HI) levels 1–6 based on estimated free hemoglobin concentration: HI 1 (0–50 mg/dL), HI 2 (51–115 mg/dL), HI 3 (116–200 mg/dL), HI 4 (201–300 mg/dL), HI 5 (301–400 mg/dL), and HI 6 (≥401 mg/dL). At our institution, potassium results are reported for specimens with HI 1–5 and suppressed for specimens with HI 6. In the central laboratory, the Vitros 7600 categorizes hemolysis according to free hemoglobin concentration, and potassium results are not reported when free hemoglobin exceeds 500 mg/dL because of unacceptable analytical interference. Thus, the reporting thresholds for potassium are >401 mg/dL free hemoglobin for the GEM Premier 7000 and >500 mg/dL for the Vitros 7600. Repeat

testing was defined as a central laboratory potassium result obtained from the same patient within 2 hours of the initial GEM 7000 measurement. Hemolysis status was assessed using the GEM hemolysis index, with subgroup analysis focused on hemolyzed samples (HI 3–6). Test utilization was evaluated by calculating the proportion of POCT results repeated in the central laboratory overall and among hemolyzed specimens. Current utilization patterns were compared with previously collected baseline data (January–April) from multiple units, including TCH WTEC, PW NICU, and TCH WTNICU, to assess changes in repeat testing practices. Descriptive statistics were used to summarize test volumes and repeat rates.

Results

A total of 673 potassium tests were performed during the one-month study period using the GEM Premier 7000 point-of-care analyzer across multiple clinical units, including ORCV, CVICU, 24ACHD, and PICU, with a smaller contribution from WTEC. Of these, 86 tests (12.78%) had a corresponding repeat potassium measurement performed in the central laboratory using the Vitros 7600 analyzer within the defined 2-hour timeframe.

Hemolysis was common in the subset of samples preselected for electrolyte measurement, with 247 of 673 samples (36.7%) categorized as hemolyzed based on hemolysis index (HI) levels 3–6. In comparison, hemolysis was relatively uncommon in the overall blood gas specimen population (~1%), although higher rates were observed in the emergency department (up to 8.2%), as reported in our prior analysis [6]

Despite relatively high prevalence, repeat testing among hemolyzed samples was infrequent. Only 20 of the 247 hemolyzed samples (8.10%) were repeated in the central laboratory, while the majority, 227 samples (91.90%), were not repeated. These findings indicate that most hemolyzed POCT potassium results did not trigger confirmatory testing.

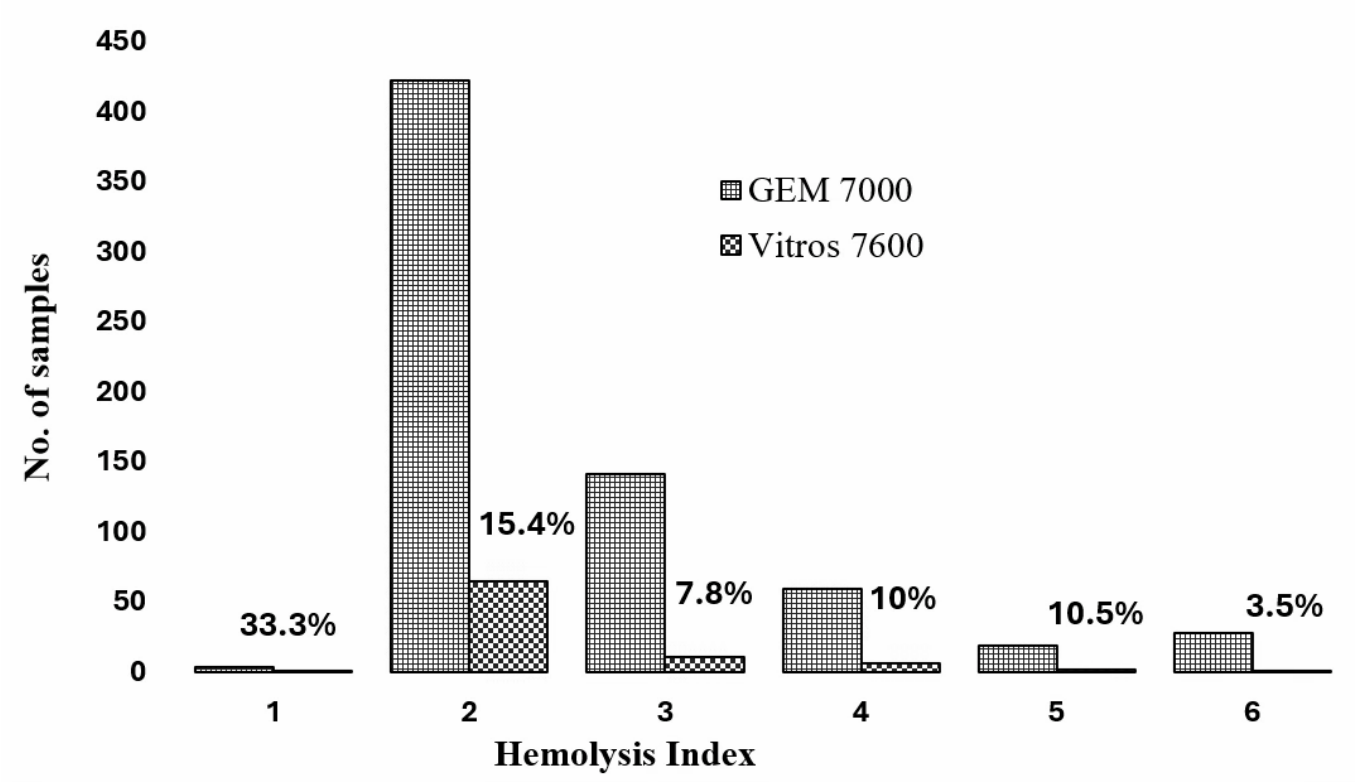
When stratified by clinical unit, variability in repeat testing rates was observed. The average reflex rate was highest in PICU (22.5%), followed by CVICU (13.6%) and OR (8.2%), indicating differences in utilization practices across care settings (Table1).

Table 1: Reflexed testing rates varied across units in relation to hemolysis index levels. Values represent the proportion of samples reflexed at each hemolysis level, with averages shown for each unit.

CVICU (n=117)		OR (n=350)		PICU (n=142)	
Hemolysis Index	**Reflexed Percentage (%)**	**Hemolysis Index**	**Reflexed Percentage (%)**	**Hemolysis Index**	**Reflexed Percentage (%)**
1	n/a	1	0	1	n/a
2	20	2	7	2	31
3	7	3	9	3	6
4	0	4	13	4	11
5	14	5	11	5	0
6	0	6	10	6	0
Avg	**13.6%**	**Avg**	**8.2%**	**Avg**	**22.5%**

Analysis by hemolysis index demonstrated that repeat testing was not uniformly distributed across HI levels. While moderate hemolysis categories (HI 2–4) accounted for the majority of repeated tests, relatively higher reflex percentages were observed in lower HI categories (HI 1–2) in some units, likely influenced by smaller sample sizes and case-specific clinical decision-making (Figure 1).

Figure 1: showing the number of total and reflexed samples across hemolysis index levels (1–6). Percent values indicate the proportion of samples reflexed at each hemolysis index relative to the total samples in that category.



Comparison between the two periods demonstrates a marked reduction in repeat testing for potassium in hemolyzed samples, decreasing from 95.05% in the baseline period to 8.10% in the current study. Additionally, overall central laboratory utilization remained low, with only 12.78% of all POCT potassium tests requiring repeat measurement. Together, these findings highlight a substantial shift in test utilization patterns following implementation of POCT across diverse clinical units.

Discussion

This study builds directly on our prior work [6] evaluating the analytical performance of the GEM Premier 7000 in a pediatric hospital setting. In that study, we demonstrated excellent agreement between GEM 7000 and central laboratory potassium measurements, along with strong concordance in hemolysis detection across platforms. While those findings established analytical reliability, the current study extends this work by examining how that reliability translates into real-world clinical practice and test utilization.

Prior study showed baseline data collected over a four-month period (January–April) from TCH WTEC, PW NICU, and TCH WTNICU identified 101 hemolyzed samples (HI 3–6), of which 96 (95.05%) underwent repeat testing in the central laboratory [6], reflecting a practice of near-universal confirmatory testing prior to POCT implementation. In the current study, repeat testing of hemolyzed samples decreased substantially from 95.05% during the baseline period to 8.10% following POCT implementation, representing a major shift from routine confirmatory testing toward a more selective and clinically driven utilization strategy.

Importantly, this change occurred despite a relatively high prevalence of hemolysis (36.7%), indicating that hemolyzed POCT potassium results were increasingly accepted without routine central laboratory confirmation.

Notably, analysis at the unit level revealed variability in repeat testing practices, with average reflex rates of 22.5% in PICU, 13.6% in CVICU, and 8.2% in OR. This heterogeneity suggests that, despite overall improvements in utilization, local clinical workflows, provider preferences, and patient acuity continue to influence repeat testing behavior. The higher reflex rate observed in PICU may reflect a more conservative approach in critically ill patients, whereas lower rates in OR settings may be driven by intraoperative workflow constraints and reliance on rapid decision-making.

This shift likely reflects increased institutional confidence in the GEM 7000 platform, particularly its ability to detect hemolysis in whole blood samples—an area that has historically limited reliance on POCT potassium testing. As previously demonstrated, the GEM 7000 provides reliable hemolysis detection with strong agreement to central laboratory systems and excellent potassium correlation. The current findings suggest that once this analytical confidence is established, it can meaningfully influence clinical behavior and reduce reliance on

redundant testing.

In addition, evaluation across hemolysis index levels demonstrated no consistent linear relationship between degree of hemolysis and likelihood of repeat testing. While moderate hemolysis (HI 2–4) accounted for the majority of repeat testing, relatively higher reflex rates were observed at lower hemolysis indices (HI 1–2) in some units. This pattern is likely influenced by smaller sample sizes in these categories and may reflect sporadic or case-specific clinical decision-making rather than a systematic response to hemolysis severity. These findings suggest that repeat testing behavior is not solely driven by hemolysis index but may be influenced by additional clinical factors such as patient condition, unexpected potassium values, or provider caution in borderline cases.

From an operational perspective, the observed reduction in repeat testing has important implications for laboratory efficiency and resource utilization. Unnecessary repeat testing contributes to increased workload, delayed turnaround times, and inefficient use of laboratory resources. In contrast, the low overall repeat rate observed in this study (12.78%) supports the role of POCT as a primary testing modality for potassium in acute care settings, with central laboratory testing reserved for selected cases where confirmation is clinically indicated. These findings are particularly relevant in pediatric and critical care environments, where minimizing blood draws and ensuring rapid availability of results are essential. Reducing unnecessary repeat testing may decrease patient discomfort, limit iatrogenic blood loss, and support more timely clinical decision-making. In this context, improved test utilization represents not only an operational benefit but also a meaningful patient-centered outcome.

The contrast between baseline and current data also highlights the importance of implementation maturity. The baseline cohort, derived from NICU-based units, demonstrated a highly conservative testing approach with near-universal repeat rates. In comparison, the current cohort, spanning multiple intensive care units, reflects a more integrated and confident use of POCT. This suggests that beyond analytical validation, factors such as user familiarity, ongoing training, and institutional trust play a critical role in shaping test utilization practices.

Taken together, these findings indicate that optimization of potassium test utilization depends not only on analytical performance, but also on appropriate interpretation and clinical use of whole blood hemolysis indices. Unit-specific variation and non-linear trends across hemolysis indices underscore the need for targeted education, standardized protocols, and ongoing monitoring to ensure consistent and appropriate use of POCT across settings.

This study has several limitations. It is a single-center retrospective analysis, and differences in patient populations and clinical units between the baseline and current cohorts may influence repeat testing behavior. Additionally, clinical outcomes were not directly assessed, and the appropriateness of individual repeat testing decisions was not evaluated. Therefore, while reduced repeat testing suggests improved efficiency, its direct impact on patient outcomes requires further study.

In summary, this study demonstrates that implementation of GEM 7000 POCT with hemolysis indices, supported by prior analytical validation, can lead to significant improvements in potassium test utilization. Together with our previous findings, these results illustrate a progression from analytical validation to operational optimization, highlighting the potential of POCT hemolysis indices to improve both laboratory efficiency and clinical workflow.

Conclusion

Implementation of point-of-care whole blood hemolysis indices improved reliance on POCT potassium results and reduced redundant potassium testing in the central laboratory, particularly for hemolyzed samples. This real-world analysis demonstrates that POCT can streamline laboratory workflows and optimize resource use while supporting timely clinical decision-making. Further studies are needed to evaluate the impact of these changes on clinical outcomes and to establish standardized approaches for managing hemolyzed POCT results.

CRediT authorship contribution statement

Nazmin Bithi: Visualization, Investigation, Formal analysis, Writing – original draft. Earnest J.P. Daniel: Writing – review & editing, Formal analysis. Michael D. Dowlin: Data Curation. Brandy Recio: Data Curation. Sridevi Devaraj:

Conceptualization

Writing – review & editing, Project administration.

Acknowledgments

Nazmin Bithi and Earnest JP Daniel were supported by the Ching-Nan Ou Endowment in Clinical Chemistry.

Funding

None.

Data Availability

Will be sent on request.

References

1. Mohamed, Taha S. Mostafa. “In Vitro Sample Hemolysis in Chemistry Laboratories and Its Challenges.” Journal of Medical Science, Biology, and Chemistry 2, no. 2 (2025): 186-191.
2. Saleh-Anaraki, Kimia, Anjuli Jain, Christopher S. Wilcox, and Negiin Pourafshar. “Pseudohyperkalemia: three cases and a review of literature.” The American Journal of Medicine 135, no. 7 (2022): e150-e154.
3. Schlüter, Kathrin, and Janne Cadamuro. “Erroneous potassium results: preanalytical causes, detection,

and corrective actions." Critical Reviews in Clinical Laboratory Sciences 60, no. 6 (2023): 442-465.
4. Toffaletti, John, and Alan HB Wu. "Correcting potassium concentrations measured in hemolyzed serum, plasma, and whole blood samples." The Journal of Applied Laboratory Medicine 10, no. 5 (2025): 1327-1332.
5. Dietzen, Dennis J., Ronald Jackups, and Mark A. Zaydman. "Clinical implications of inaccurate potassium determination in hemolyzed pediatric blood specimens." Clinica Chimica Acta 557 (2024): 117862.
6. Ibrahim, Ridwan B., Nazmin Bithi, Charlene Hernandez, Hanna Uhrova, Brandy D. Recio, and Sridevi Devaraj. "Hemolysis detection using the GEM 7000 at the point of care in a pediatric hospital setting: does it affect outcomes?." Clinical Chemistry and Laboratory Medicine (CCLM) 63, no. 12 (2025): e261-e265.
7. Mohamed, Taha S. Mostafa. "In Vitro Sample Hemolysis in Chemistry Laboratories and Its Challenges." Journal of Medical Science, Biology, and Chemistry 2, no. 2 (2025): 186-191.
8. von Petersdorff-Campen, Kai, and Marianne Schmid Daners. "Hemolysis testing in vitro: a review of challenges and potential improvements." ASAIO journal 68, no. 1 (2022): 3-13.
9. Balasubramanian, Shankar, Emily J. McDowell, Erving T. Laryea, Gert Blankenstein, Prasad VA Pamidi, Anne M. Winkler, and James H. Nichols. "Novel in-line hemolysis detection on a blood gas analyzer and impact on whole blood potassium results." Clinical Chemistry 70, no. 12 (2024): 1485-1493.
10. Wilson, Amanda, Maranna Sweeney, P. L. Mark Lynch, and Maurice J. O'Kane. "Hemolysis rates in whole blood samples for blood gas/electrolyte analysis by point-of-care testing." Journal of Applied Laboratory Medicine 3, no. 1 (2018): 144-145.
11. DiToro, Daniel F., Michael J. Conrad, and Petr Jarolim. "Hemolysis index and potassium reporting: contextual evidence-based reporting criteria." American Journal of Clinical Pathology 157, no. 6 (2022): 809-813.
12. Khodorkovsky, Boris, Bartholomew Cambria, Martin Lesser, and Barry Hahn. "Do hemolyzed potassium specimens need to be repeated?." The Journal of Emergency Medicine 47, no. 3 (2014): 313-317.
13. Khattak, Zoia E., Husam El Sharu, and Beenish S. Bhutta. "Overview on ordering and evaluation of laboratory tests." In StatPearls [Internet]. StatPearls Publishing, 2023.
14. Elrobaa, Islam H., Keebat Khan, and Eslam Mohamed. "The role of point-of-care testing to improve acute care and health care services." Cureus 16, no. 3 (2024).
15. Pighi, Laura, Gian Luca Salvagno, Filippo Marcazzan, Mariateresa Rizza, and Giuseppe Lippi. "Improved potassium reliability in whole blood through hemolysis detection on the novel GEM Premier 7000 blood gas analyzer." Journal of Medical Biochemistry 44, no. 4 (2025): 915.

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:

ifcc
International Federation of Clinical Chemistry and Laboratory Medicine

www.ifcc.org