Advances in Clinical and Experimental Medicine

MONTHLY ISSN 1899-5276 (PRINT) ISSN 2451-2680 (ONLINE) advances.umw.edu.pl

2026, Vol. 35, No. 6 (June)

Impact Factor (IF) – 2.5
Ministry of Science and Higher Education – 70 pts
Index Copernicus (ICV) – 161.00 pts

Wroclaw Medical University

Advances
in Clinical and Experimental
Medicine

Advances
in Clinical and Experimental Medicine

ISSN 1899-5276 (PRINT) ISSN 2451-2680 (ONLINE) advances.umw.edu.pl

MONTHLY 2026
Vol. 35, No. 6
(June)

Advances in Clinical and Experimental Medicine (*Adv Clin Exp Med*) publishes high-quality original articles, research-in-progress, research letters and systematic reviews and meta-analyses of recognized scientists that deal with all clinical and experimental medicine.

Editorial Office
ul. Marcinkowskiego 2–6
50-368 Wrocław, Poland
Tel.: +48 71 784 12 05
E-mail: acem.journal@umw.edu.pl

Publisher
Wroclaw Medical University
Wybrzeże L. Pasteura 1
50-367 Wrocław, Poland

Online edition is the original version of the journal

Editorial Policy

Advances in Clinical and Experimental Medicine (Adv Clin Exp Med) is an independent multidisciplinary forum for exchange of scientific and clinical information, publishing original research and news encompassing all aspects of medicine, including molecular biology, biochemistry, genetics, biotechnology and other areas. During the review process, the Editorial Board conforms to the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication" approved by the International Committee of Medical Journal Editors (www.ICMJE.org). The journal publishes (in English only) original papers and reviews. Short works considered original, novel and significant are given priority. Experimental studies must include a statement that the experimental protocol and informed consent procedure were in compliance with the Helsinki Convention and were approved by an ethics committee.

For all subscription-related queries please contact our Editorial Office: acem.journal@umw.edu.pl
For more information visit the journal's website: advances.umw.edu.pl

Pursuant to the ordinance of the Rector of Wroclaw Medical University No. 37/XVI R/2024, from March 1, 2024, authors are required to pay a fee for each manuscript accepted for publication in the journal Advances in Clinical and Experimental Medicine. The fee amounts to 1600 EUR for all types of papers.

Indexed in: MEDLINE, Science Citation Index Expanded, Journal Citation Reports/Science Edition, Scopus, EMBASE/Excerpta Medica, Ulrich's™ International Periodicals Directory, Index Copernicus

Typographic design: Piotr Gil, Monika Kolęda
DTP: Wydawnictwo UMW
Cover: Monika Kolęda
Printing and binding: Agencja Wydawnicza "ARGI" s.c.

Advances
in Clinical and Experimental Medicine

MONTHLY 2026, Vol. 35, No. 6 (June)

ISSN 1899-5276 (PRINT)
ISSN 2451-2680 (ONLINE)
advances.umw.edu.pl

Contents

Reviews

Academic mediator-driven research translation: The concept for making possible implementation

Sajee Sattayut[1,A–F], Patcharawan Srisilapanan[2,E,F], Piyachat Patcharanuchat[1,3,E,F]

[1] Lasers in Dentistry Research Group, Faculty of Dentistry, Khon Kaen University, Thailand
[2] School of Dentistry, Phayao University, Thailand
[3] Department of Preventive Dentistry, Faculty of Dentistry, Khon Kaen University, Thailand

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):949–955

Address for correspondence
Sajee Sattayut
E-mail: sajee@kku.ac.th

Funding sources
None declared

Conflict of interest
None declared

Received on November 9, 2025
Reviewed on March 11, 2026
Accepted on April 11, 2026

Published online on April 21, 2026

Abstract

Translation research into practice remains one of the critical yet underdeveloped aspects of scientific progress. This editorial explores how academic mediation can transform fragmented research activities into sustainable translational ecosystems. Drawing from Thailand's experience in laser dentistry, the paper illustrates a progression from individual projects to collective research networks and the institutionalization of the Hub of Knowledge in Orofacial Laserology (HKOL). The model demonstrates how innovation can move beyond diffusion and normalization to achieve implementation grounded in education and community relevance through structured collaboration, standardized protocols, and academic stewardship. By addressing theoretical gaps and aligning practice with governance, the academic mediator-driven model offers an adaptable framework for integrating research, clinical service, and policy. It underscores that effective translation is not a spontaneous outcome of evidence but the result of deliberate structure, shared accountability, and long-term academic continuity.

Key words: oral health, development, laser therapy, evidence-based practice, translational science

Cite as
Sattayut S, Srisilapanan P, Patcharanuchat P. Academic mediator-driven research translation: The concept for making possible implementation. *Adv Clin Exp Med*. 2026;35(6):949–955. doi:10.17219/acem/220569

DOI
10.17219/acem/220569

Highlights

- Translational research in healthcare requires structured academic mediation to bridge the gap between scientific evidence and real-world clinical practice.
- The Hub of Knowledge in Orofacial Laserology (HKOL) model demonstrates how collaborative research networks can transform fragmented studies into sustainable translational ecosystems.
- Standardized protocols, education, and academic leadership are key drivers for successful implementation of innovation in fields such as laser dentistry.
- An academic mediator-driven approach enables effective integration of research, clinical practice, and health policy, ensuring long-term impact and scalability.

Introduction

All academic knowledge and innovation in medicine ultimately aim to improve clinical practice, often described as the pathway from bench to bedside. However, evidence indicates a substantial delay between discovery and adoption: a review of 23 studies from the 1980s to the 2000s revealed an average lag of approx. 17 years before research findings achieved widespread clinical implementation.[1] Meanwhile, the linkage between biomedical patents and scientific papers has expanded exponentially over the past 37 years, doubling every 2.9 years, particularly among biotechnology and U.S. Food and Drug Administration (FDA)-related drug patents.[2]

Despite this growing scientific output, the persistent gap between research and patient benefit remains evident, as reflected in recent systematic reviews published between 2023 and 2025.[3–5] Worldwide, research institutions produce an unprecedented volume of studies; however, only a small proportion of these findings is translated into improvements in clinical practice, public health policy, or community wellbeing.[6,7] The current challenge is no longer how to create knowledge but rather how to make that knowledge meaningful.

This issue is particularly critical in healthcare systems of low- and middle-income countries, where innovation often relies on imported technologies and external evidence. Despite global investments in research, implementation is often hampered by a limited research and development ecosystem, including inadequate infrastructure, fragmented governance, and a lack of local ownership. As Alonge et al.[6] noted, advancing equity in implementation research requires adapting evidence to context and transforming the power relations that determine whose knowledge is valued and whose health truly improves.

Even in high-income countries like the UK and Australia, integrating research, education, and healthcare delivery remains a challenge. In the UK, initiatives such as Academic Health Science Centers, Networks, and Collaborations (AHSCs) aim to promote innovation in the National Health Service (NHS). Similarly, Australia's National Health and Medical Research Council (NHMRC) accredits centers to enhance health research translation. Based on reviews of 41 projects or centers under those organizations on the website by Robinson et al.,[8] barriers such as fragmented governance, short funding cycles, uneven public involvement, and limited long-term evaluation persist. Infrastructure alone has not ensured integration; strong leadership and cross-sector collaboration remain crucial.

Recent editorials in orthopedics and sports medicine have highlighted the concept of evidence-based research (EBR), emphasizing that new studies must rely systematically on prior evidence and ensure the practical implementation of findings rather than merely contributing to publication volume. Prill et al.[9] argue that research achieves value only when it is transparently justified, methodologically rigorous, and integrated into clinical decision-making and guideline development. This perspective reframes implementation not as a downstream activity but as an essential dimension of responsible evidence generation. Within this broader evidence-based discourse, academic mediation can be understood as an institutional mechanism that operationalizes EBR principles across healthcare systems.

Over the past decade, several theoretical frameworks, such as Diffusion of Innovations,[10] Normalization Process Theory,[5] and Implementation Science,[11,12] have attempted to explain and assess the causes of delayed translation of research to practice in healthcare services. While each provides valuable insights, their practical impact has varied. Translation often diverges from the neat patterns proposed by these theories, unfolding in complex, context-specific ways influenced by multiple levels of relationships, including connections among academic institutions, hospitals, and grading agencies, as well as administration, trust, and social and organizational values.

In this context, Thailand's experience offers a unique perspective on the role of translation and implementation under academic mediation. Instead of keeping research confined within universities, the Thai model positions academia as a bridge connecting clinical researchers and communities through micro- and meso-level collaboration. In the field of orofacial laserology, this approach blends isolated clinical trials with translational research and implementation, along with workforce capacity building through networks.

Objectives

This editorial addresses the role of academic mediators in research translation, focusing on the achievements of the Lasers in Dentistry Group at Khon Kaen University and the Hub of Knowledge in Orofacial Laserology (HKOL) in Thailand. In this editorial, academic mediation refers to the institutional function of universities in aligning evidence generation with clinical implementation and policy coordination. It begins by examining the reasons behind the frequent failures of research translation on a global scale. The editorial then explores how the intentional creation of research networks and transfer mechanisms, grounded in education and community partnership, can convert knowledge into sustainable practices. The Thai experience illustrates that when innovations are directly transferred to local practices through academic networks rather than through a top-down system, they can spread and be sustained, effectively transforming research into practice. Additionally, a content analysis is presented alongside relevant theories.

From research to reality: Why translation often fails and when it succeeds

Over decades of scientific progress, one paradox remains: while research output has grown exponentially, only a small portion of this knowledge is ever put into practice. Numerous studies, ranging from the study by Grimshaw et al.[11] in 2012 to Williams et al.[5] in 2023, consistently show that most research remains unutilized, trapped within the academic cycle of publication rather than being translated into real-world applications. In healthcare, this gap between research and practice has become one of the most enduring challenges of modern science. Translation fails not because of a lack of evidence, but because of a lack of mediation. The Thai model of orofacial laserology demonstrates how networks of academic mediation can operationalize theory into practice. According to previous reviews and analyses, several recurring reasons explain why translation frequently stalls.

Linear diffusion myth

The classical Diffusion of Innovations model suggests that new ideas spread naturally once they are proven effective.[13] However, Greenhalgh and Papoutsi[14] argue that this linear and individual-focused model overlooks the complex and adaptive nature of real health systems. In these systems, social, political, and institutional factors play a crucial role in determining whether innovations are sustainable or fade away. They emphasize that sustainability depends not only on the spread of an innovation but also on ongoing sense-making, local adaptation, and alignment with the values and resources of the system – processes that the traditional diffusion curve cannot capture.

Dissemination but not implementation

Merely publishing or presenting findings does not ensure application. Grimshaw et al.[11] demonstrated that the distance between what is known and what is done persists because academic channels, such as papers, conferences, and presentations, rarely connect with practitioners' daily realities.

Organizational and cultural barriers

System-level inertia often hinders the translation of evidence. Abu-Odah et al.[3] conducted systematic reviews that identified barriers such as limited leadership incentives, scarce resources, and a lack of structured feedback loops as major reasons why evidence fails to influence policy or practice. These results illustrate the common global situation in healthcare, where the research and service sectors often operate as parallel universes.

The challenge of using theories without implementation

Ironically, even implementation science, the discipline created to close the gap, has developed its own gap. Westerlund et al.[7] described this as the "implementation paradox," where new frameworks proliferate but remain unadopted by practitioners. A similar pattern is evident with Normalization Process Theory (NPT): although widely cited, few studies have applied its core principles in real-world clinical settings. Williams et al.[5] reviewed trials referencing NPT and found that only a small number operationalized its mechanisms to support genuine normalization in practice.

Malalignment of scientific advancements and clinical application

Despite rapid scientific advancements, the path from discovery to patient benefit remains misaligned. Fernandez-Moure[15] identified key obstacles, such as reduced protected time for clinician-scientists, academic incentives favoring publication over application, and significant funding shortages between initial discovery and implementation.

These issues delay innovation and hinder promising technologies from gaining the necessary support for regulatory approval and real-world use. Thus, the challenge lies not in a lack of knowledge, but in fragmented systems that fail to support its application.

Unsustainable collaborations

Without long-term structures, collaborative projects dissolve once funding ends. Systematic reviews of the websites

of research centers and projects showed that many university–health service partnerships fade after the project phase, lacking governance mechanisms to sustain shared learning and accountability.[8]

What enables translation and implementation

Across the literature, 5 enabling conditions consistently appear[3,5,7,8,11,14,15]:

1. Co-production of knowledge: Engaging end-users such as healthcare providers, patients, and stakeholders early in research design ensures that questions and outcomes are meaningful beyond academia.
2. Academic mediation: Universities and research networks are trusted intermediaries that interpret findings, train users, and monitor outcomes.
3. Feedback integration: Real-world data loop into research, closing the cycle between evidence and improvement.
4. Incentivized networks: Translation thrives when collaboration, not competition, is rewarded.
5. Adaptive education: Embedding translational principles within curricula cultivates practitioners who can generate and apply evidence.

When these elements align, research ceases to be an isolated endeavor. It becomes a living, self-renewing ecosystem. The following lessons were learned from the Lasers in Dentistry Research Group (LDRG) in Thailand. Despite facing ongoing challenges, these collective failures highlight how successful translation can occur. The translation of laser use for oral health in community hospitals becomes achievable when evidence effectively functions within a real network of equality. Academia takes on a purposeful role as the translator, trainer, and guardian of innovation. This is where the Thai model offers its most valuable insights.

From individual discovery to collective networks

Thailand's experience with orofacial laserology vividly illustrates these lessons. The development of this field signifies more than just technological advancement; it embodies a continuous journey of exploration, adaptation, and integration that connects academic innovation with public health priorities.

In 1993, when Khon Kaen University's Faculty of Dentistry acquired its first dental laser, it represented more than just an equipment purchase; it marked the beginning of an intellectual movement. By 1998, the first Thai specialist had completed a PhD in laser dentistry, establishing domestic expertise and reducing reliance on imported knowledge. By 2001, laser courses were formally integrated into both undergraduate and continuing education curricula, cultivating foundational literacy across dental disciplines. An expansion in 2007, supported by national and international partners, positioned Thai initiatives within a broader professional community.

The LDRG at the Khon Kaen University, founded in 2011, formalized the transition from isolated projects to a coherent research network. Among the various approaches to conducting research and innovation in dentistry, the group's early laboratory and clinical studies clarified how low-intensity photobiomodulation and diode laser photocoagulation could be adapted for mucosal surgery, soft-tissue welding, and wound control.[16–18] These techniques were accepted and adopted by tertiary hospitals via participatory action research due to the demand to treat oral potentially malignant disorders.[19]

By 2017, collaboration expanded beyond tertiary hospitals to include local health services in rural areas. This shift was driven by the clinical demand for safe, minimally invasive procedures and effective hemostatic support, all in line with the goal of promoting oral health. Between 2020 and 2022, a prototype low-intensity diode laser for primary care was co-developed with local clinicians and engineers, with support from the National Research Council of Thailand (NRCT). The device, which operates at a compact wavelength of 810–830 nm, offers intensity ranges from 0.2 W to 1.0 W, optimized for wound healing, pain control, and blood clot initiation. It was specifically designed for health-promoting hospitals to address the needs of older adults with systemic health risks while aligning with the demands and utilization practices of dental healthcare providers.

This initiative demonstrated that successful translation requires not only a device but also an ecosystem of learning. Community hospitals became both service and learning centers, forwarding anonymized data to a PhD candidate, including a multicenter hemostasis study by Tanya et al. in 2025.[20] These clinical translational studies have confirmed the safety and effectiveness of laser-mediated hemostasis and photobiomodulation therapy, which promote healing and alleviate pain for older patients. The results consistently showed that socket hemostasis could be achieved within minutes, which improved treatment acceptance, particularly among patients on antithrombotic therapy. Additionally, there was a positive response from dental healthcare providers in local hospitals regarding the technological aspects of these advancements.

Academic mediation and the Hub Model

The next transformation came with the recognition that sustainability required structure.

In 2024, the HKOL, supported by the NRCT, was established to integrate research, education, and clinical service into a single translational system.

The need for this formal structure became evident when the first primary-care network, Bua Ngoen Primary Hospital (Khon Kaen, Thailand), successfully translated the routine use of lasers for oral healthcare in primary hospitals to 7 neighboring health-promoting hospitals by organizing local workshops and peer-to-peer learning. This demonstrated that knowledge transfer could truly propagate through academic mediation in practice rather than theory.

As expansion reached a nationwide scale, new challenges arose. According to Thailand's public procurement regulations, acquiring laser devices for primary and secondary hospitals requires clear and standardized specifications. The HKOL took on an official role in defining academic specifications for each level of healthcare to ensure that quality, safety, and educational alignment were maintained. These specifications align well with the real-world experiences of the Lasers in Dentistry networks at Khon Kaen University and the HKOL, which have been utilized for at least 10 years.

The HKOL established standardized device categories: low-intensity presets for primary hospitals, adjustable diode systems for secondary facilities, and CO_2 or Er:YAG lasers for tertiary centers. These were all implemented under unified safety protocols, operator certification, and academic endorsement. Additionally, the HKOL coordinated postgraduate research, national workshops, and collaborative data collection in community hospitals, effectively transforming research translation into an organized, self-sustaining system.

The HKOL embodies the principle of academic mediation: knowledge transfer as stewardship, not diffusion. Evidence becomes education; education becomes service; and service feeds new evidence. The model's success lies in its continuity, supported not by transient grants but by a shared academic identity across institutions. It is a living network that keeps translation alive in lifelong education.

From local innovation to global learning

Thailand's trajectory in orofacial laserology illustrates what global literature repeatedly calls for but seldom achieves: a functioning translational ecosystem.

It demonstrates that innovation can originate from a middle-income setting, provided the process is relationship-driven, ethically grounded, and academically mediated. Local research became globally relevant through continuous collaboration linking the LDRG, HKOL, and international partners such as the World Federation for Laser Dentistry (WFLD). What began as a clinical response to limited surgical options for older adults evolved into a community-based translational model that others can learn from. In this model, transfer is not a one-time event but an ongoing conversation among researchers, practitioners, and patients.

From barriers to bridges

The global literature highlights the reasons for translation failures and suggests that success can emerge when these failures are intentionally addressed. While frameworks like diffusion, normalization, and implementation science each offer insights into the translational process, none guarantees sustained progress without a mediating structure to integrate them.

In Thailand, this structure is achieved through academic mediation, where universities serve as both sources of innovation and custodians of translation. By turning research networks into engines of education and collaboration, the orofacial laserology community demonstrates that theory is only durable when linked to practice, which, in turn, must be grounded in evidence.

The analytical synthesis (Table 1 [3,5,7,8,12–14]) summarizes how this network-based academic mediation model overcomes theoretical limitations by transforming diffusion into direction, embedding into endurance, and implementation into institutionalization.

Discussion

Academic mediation does not introduce a new theoretical framework. It serves as an operational layer that integrates existing implementation frameworks. While diffusion theory explains how innovations spread, and normalization process theory outlines the integration of practices into routine, neither alone ensures continuity across institutions.

Academic mediation institutionalizes feedback loops, aligns educational practices, and coordinates policies, functioning as the infrastructure that transforms theoretical insights into sustainable practices. Thus, it represents the institutional architecture for implementing evidence-based practices. The synthesis presented in Table 1 demonstrates how the academic mediator-driven translation model transforms theoretical frameworks into practical applications. Instead of dismissing diffusion, normalization, or implementation science, our approach operationalizes these frameworks by providing institutional continuity and structured feedback mechanisms linking evidence generation, education, and clinical service. This transformation was intentional and reflected the development philosophy of His Majesty King Bhumibol Adulyadej (Rama IX), who emphasized that sustainable progress must begin with "Understand, Access, and Develop" and should be driven by "development from within."[21]

The process highlighted the importance of contextual relevance by prioritizing an accurate understanding of the networks prior to implementation. It encouraged collaboration among micro- and meso- levels of universities, community hospitals, and practitioners by ensuring resource access. Additionally, fostering development within

Table 1. Addressing theoretical gaps in research translation through the Thai model of academic mediation and laserology networks

Theoretical framework	Typical limitations or gaps	How the Thai academic mediation model addressed these gaps	Outcome/evidence of success
Diffusion of Innovations[13]	Assumes innovation spreads naturally once proven effective; neglects system complexity.	Replaced passive diffusion with guided dissemination through structured academic mediation.	Controlled diffusion across healthcare tiers; unified safety standards; consistent outcomes.
Normalization Process Theory[5]	Describes embedding but rarely operationalized; lacks a mechanism for reinforcement.	Embedding through the Hub of Knowledge in Orofacial Laserology (HKOL), thus institutionalizing education, research, and practice.	Sustainable integration via certification, workshops, and ongoing lifelong education.
Implementation Science Paradox[7]	Frameworks were produced, but minimal practitioner adoption.	Practice-led research: community clinicians co-research and feed data into academia.	Ownership of innovation by end-users; co-production closes the 'gap within the gap'.
Organizational barriers[3]	Fragmented policy, limited incentives, and a lack of feedback loops.	Created continuous feedback integration from field practice into research refinement.	Academic empowerment and guardian of the primary care team for reflection on the policy maker.
Academic reward systems[14]	Rewards novelty and publication, not application; undervalues teamwork.	Reframed success to include translational impact in postgraduate evaluation.	Culture shift: young PhD pursue applied innovations.
Temporal instability of collaborations[8]	Partnerships collapse after funding ends.	Anchored translation within university governance: HKOL as a permanent platform.	Longevity beyond projects; nationwide multi-institutional collaboration.
Limited community co-production[12]	Communities are seen as recipients, not co-producers of innovation.	Engaged primary care staff and older adults in co-designing a low-intensity laser prototype and recognized them as part of the learning resources.	Higher acceptance, better satisfaction, and measurable oral-health improvement.

the local organization, in this case, enhanced local capabilities and ownership. Through this approach, academic mediation transcended its role as merely a means of translation; it became a form of stewardship in which knowledge flows between academia and the community, continuously sustained through education and shared responsibility.

The HKOL and its national collaborations illustrate the evolution of the field. Supported by the NRCT, the HKOL transformed translation from a mere project into a comprehensive system that integrates innovation, governance, and compassion. This model from Thailand demonstrates that effective translation research relies not just on scientific validity but also on moral clarity: progress should be rooted in empathy, participation, and internal growth.

Ultimately, the Thai experience shows that compassionate science serves as the true bridge between theory and practice. When universities act as mediators that listen before acting, reach out before prescribing solutions, and empower before directing.

The transferability of this operational model to other healthcare systems relies on several key contextual factors. First, a minimum level of academic infrastructure is necessary, which includes high research capacity and a strong institutional commitment to long-term collaboration. Second, funding mechanisms should prioritize continuity over short-term project cycles. Third, clear regulations, especially regarding procurement and professional certification, are important for standardization across healthcare tiers. Lastly, building relational trust among universities, healthcare personnel, and related administrative organizations is essential. Without these enabling factors, academic mediation may remain an aspirational goal rather than become operational. While the Thai context is one example, the fundamental principle of embedding implementation within institutional governance can be applied to various systems with appropriate contextual adjustments.

Conclusions

The development of research networks in orofacial laserology in Thailand is a good example that successful translation depends on institutional structure rather than technological novelty alone. The Thai model operationalized innovation through a multi-level system linking university-based research, primary care services, and postgraduate education within a coordinated governance framework. Through academic mediation, universities functioned not merely as producers of evidence but as institutional stewards aligning research generation, professional training, and community implementation.

This experience suggests 3 broader implications for research translation:

1. Implementation should be embedded within institutional governance rather than treated as a downstream activity following evidence generation.
2. Sustainability requires permanent networks and standardized protocols that transcend project-based funding cycles and support continuity across healthcare levels.
3. Co-creation with practitioners, supported by structured feedback loops, is essential to transform evidence into routine clinical practice.

While contextual conditions influence transferability, the underlying principle remains consistent: research

achieves impact when supported by structured collaboration, accountability mechanisms, and long-term institutional commitment. Translation, therefore, is best understood not as a linear transfer of knowledge but as a sustained organizational process integrating evidence, education, and service within a continuous learning ecosystem.

Use of AI and AI-assisted technologies

AI-assisted tools were used partially for grammatical checking and for preparing the graphical abstract based on the authors' design. The tools that we used were OpenAI's ChatGPT (GPT-5.4 Thinking) and Grammarly.

ORCID iDs

Sajee Sattayut https://orcid.org/0000-0001-7111-9381
Patcharawan Srisilapanan https://orcid.org/0000-0001-9407-7452
Piyachat Patcharanuchat https://orcid.org/0009-0006-1096-9415

References

1. Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: Understanding time lags in translational research. *J R Soc Med*. 2011;104(12):510–520. doi:10.1258/jrsm.2011.110180
2. Ke Q. An analysis of the evolution of science-technology linkage in biomedicine. *J Informetrics*. 2020;14(4):101074. doi:10.1016/j.joi.2020.101074
3. Abu-Odah H, Said NB, Nair SC, et al. Identifying barriers and facilitators of translating research evidence into clinical practice: A systematic review of reviews. *Health Soc Care Community*. 2022;30(6):e3265–e3276. doi:10.1111/hsc.13898
4. Batchelor J, Hemmert C, Meulenbroeks I, et al. Factors influencing the translation of evidence into clinical practice for hospital allied health professionals in terms of the domains of behaviour change theory: A systematic review. *Eval Health Prof*. 2025;48(4):471–489. doi:10.1177/01632787241285993
5. Williams A, Lennox L, Harris M, Antonacci G. Supporting translation of research evidence into practice: The use of Normalisation Process Theory to assess and inform implementation within randomised controlled trials. A systematic review. *Implement Sci*. 2023;18(1):55. doi:10.1186/s13012-023-01311-1
6. Alonge O. How to leverage implementation research for equity in global health. *Glob Health Res Policy*. 2024;9(1):43. doi:10.1186/s41256-024-00388-5
7. Westerlund A, Nilsen P, Sundberg L. Implementation of implementation science knowledge: The research–practice gap paradox. *Worldviews Evid Based Nurs*. 2019;16(5):332–334. doi:10.1111/wvn.12403
8. Robinson T, Bailey C, Morris H, et al. Bridging the research–practice gap in healthcare: A rapid review of research translation centres in England and Australia. *Health Res Policy Sys*. 2020;18(1):117. doi:10.1186/s12961-020-00621-w
9. Prill R, Pieper D, Klugar M, Ayeni OR, Karlsson J, Lund H. Evidence-based research in orthopaedics, sports medicine and rehabilitation: Why new studies should rely on earlier work. *Knee Surg Sports Traumatol Arthrosc*. 2024;32(2):203–205. doi:10.1002/ksa.12047
10. Dearing JW, Cox JG. Diffusion of innovations theory, principles, and practice. *Health Aff (Millwood)*. 2018;37(2):183–190. doi:10.1377/hlthaff.2017.1104
11. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. *Implement Sci*. 2012;7(1):50. doi:10.1186/1748-5908-7-50
12. Finney Rutten LJ, Ridgeway JL, Griffin JM. Advancing translation of clinical research into practice and population health impact through implementation science. *Mayo Clin Proc*. 2024;99(4):665–676. doi:10.1016/j.mayocp.2023.02.005
13. Rogers EM. Diffusion of preventive innovations. *Addict Behav*. 2002;27(6):989–993. doi:10.1016/S0306-4603(02)00300-3
14. Greenhalgh T, Papoutsi C. Spreading and scaling up innovation and improvement. *BMJ*. 2019;365:l2068. doi:10.1136/bmj.l2068
15. Fernandez-Moure JS. Lost in translation: The gap in scientific advancements and clinical application. *Front Bioeng Biotechnol*. 2016;4:43. doi:10.3389/fbioe.2016.00043
16. Malakam C, Rungrojwittaya N, Sanposh N, Sattayut S. A variety of photocoagulation techniques: Haemostasis in oral soft tissue and extraction socket. *Laser Int Magaz Laser Dent*. 2016;8(4):16–20. https://www.academia.edu/93765654/A_variety_of_photo_coagulation_techniques_Haemostasis_in_oral_soft_tissue_and_extraction_socket.
17. Saenthaveesuk P, Sanjandee N, Treeratsakulchai T, Norateethan P, Sattayut S. The effect of oral tissue welding technique using diode laser. *J Dent Res*. 2013;92(Special Iss B):591 (IADR-APR). https://iadr.abstractarchives.com/abstract/iadr-apr2013-181205/theeffect-of-oral-tissue-welding-technique-using-diode-laser.
18. Sattayut S. Novel technique for using the diode laser to treat refractory erosive oral lichen planus. *Laser*. 2011;3:18–20. https://iadr.abstractarchives.com/abstract/iadr-apr2013-181205/theeffect-of-oral-tissue-welding-technique-using-diode-laser.
19. Sattayut S, Chaimusig M, Patcharanuchat P. A review of efficacy of laser therapy on complicated or non-responsive oral soft tissue lesions in elderly patients in 3 hospitals, Thailand. *Med Oral Patol Oral Cir Bucal*. 2012;17(Suppl 1):S103. doi:10.4317/medoral.17643602
20. Tanya S, Prajaneh S, Patcharanuchat P, Sattayut S. Laser-mediated hemostasis for older patients receiving routine dental treatment. *Dent J (Basel)*. 2025;13(7):315. doi:10.3390/dj13070315
21. United Nations Development Programme (UNDP). *Thailand Human Development Report 2007: Sufficiency Economy and Human Development*. Bangkok, Thailand: United Nations Development Programme (UNDP); 2007. https://hdr.undp.org/system/files/documents/thailand2007en.pdf.

Prognostic significance of ctDNA mutations in advanced HER2-positive breast cancer treated with targeted therapy: A meta-analysis

Jialin Lin[A–D], Hangcheng Xu[B], Yiran Zhou[B], Qiang Sa[B], Jiayu Wang[E], Binghe Xu[F]

Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):957–968

Address for correspondence
Binghe Xu
E-mail: xubinghebm@163.com

Funding sources
This study was supported by the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (CIFMS) (grants No. 2021-12M-1-014 and No. 2022-12M-2-002).

Conflict of interest
None declared

Received on June 4, 2025
Reviewed on August 2, 2025
Accepted on August 27, 2025

Published online on June 25, 2026

Abstract

Background. Circulating tumor DNA (ctDNA) is a promising noninvasive biomarker in advanced breast cancer. In patients with advanced *HER2*-positive disease, responses to targeted therapy vary. The prognostic significance of ctDNA mutations across different genes and treatment regimens remains to be fully clarified.

Objectives. This meta-analysis evaluated the predictive role of ctDNA mutations in guiding anti-*HER2* therapy.

Materials and methods. We searched PubMed, Embase, the Cochrane Library, and major oncology conference proceedings for studies evaluating associations between ctDNA and progression-free survival (PFS) or overall survival (OS) in breast cancer. Pooled hazard ratios (HRs) and odds ratios (ORs) were calculated using a pre-specified random-effects model. Study quality, publication bias, and robustness were assessed using the Newcastle–Ottawa Scale (NOS), funnel plots, Egger's test, and leave-one-out analyses.

Results. A total of 12 studies involving 558 patients were included. Baseline ctDNA mutation status was significantly associated with shorter PFS (HR = 1.73, 95% confidence interval (95% CI): 1.06–2.82, p = 0.03). *PIK3CA* mutations correlated with worse PFS (HR = 2.12, 95% CI: 1.43–3.15, p = 0.002), whereas *ERBB2* mutations showed no significant association. ctDNA mutations predicted poorer PFS in tyrosine kinase inhibitor (TKI)-treated patients (HR = 2.04, 95% CI: 1.29–3.24, p = 0.002), particularly in patients receiving pyrotinib (HR = 2.77, 95% CI: 1.96–3.92, p < 0.001), but not in those receiving non-pyrotinib TKIs. Worse PFS was also observed in monotherapy (HR = 2.65, 95% CI: 1.81–3.88, p < 0.001) and capecitabine-based combination regimens (HR = 2.07, 95% CI: 1.22–3.50, p = 0.007), but not in other combination therapies.

Conclusions. Circulating tumor DNA mutation status may serve as a prognostic biomarker in patients receiving *HER2*-targeted therapies. *PIK3CA* mutations were associated with worse outcomes, whereas *ERBB2* mutations showed no significant effect. The prognostic significance of ctDNA was most pronounced in patients treated with TKIs, particularly pyrotinib, and remained evident in monotherapy and capecitabine-based regimens. These findings support the potential utility of ctDNA monitoring for risk stratification and personalized management in advanced *HER2*-positive breast cancer.

Key words: *PIK3CA*, ctDNA, breast cancer, anti-*HER2* therapy, TKIs

Cite as
Lin J, Xu H, Zhou Y, Sa Q, Wang J, Xu B. Prognostic significance of ctDNA mutations in advanced HER2-positive breast cancer treated with targeted therapy: A meta-analysis. *Adv Clin Exp Med*. 2026;35(6):957–968. doi:10.17219/acem/209957

DOI
10.17219/acem/209957

Highlights

- Circulating tumor DNA (ctDNA) mutations predict poor progression-free survival in advanced *HER2*-positive breast cancer.
- *PIK3CA* mutations in ctDNA are strongly associated with worse outcomes, while *ERBB2* mutations show no prognostic impact.
- Prognostic significance of ctDNA mutations is most evident in patients treated with tyrosine kinase inhibitors, especially pyrotinib.
- ctDNA mutation profiling offers a noninvasive tool for risk stratification and personalized management in *HER2*-targeted therapy.

Introduction

HER2-positive breast cancer, characterized by *HER2* gene amplification or overexpression, leads to continuous activation of the PI3K/AKT/mTOR and RAS/MAPK signaling pathways, resulting in increased aggressiveness and poor prognosis.[1–4] The advent of *HER2*-targeted therapies has significantly improved patient outcomes; however, individual responses vary, with some patients experiencing primary or acquired resistance, leading to reduced efficacy and disease progression.[5] Current *HER2*-targeted therapies include monoclonal antibodies such as trastuzumab and pertuzumab, antibody-drug conjugates (ADCs) such as T-DM1 and T-DXd, and tyrosine kinase inhibitors (TKIs) including lapatinib, pyrotinib, and neratinib. Despite their efficacy across various clinical stages, resistance mechanisms are complex and involve *HER2* mutations, epithelial–mesenchymal transition (EMT), and alterations in the tumor microenvironment (TME).[6] Therefore, accurately predicting the efficacy of *HER2*-targeted treatments and dynamically monitoring patient responses are crucial for optimizing personalized treatment strategies.

Circulating tumor DNA (ctDNA), a noninvasive biomarker reflecting tumor genetic alterations, has garnered attention in the diagnosis, prognosis, and therapeutic monitoring of breast cancer.[7–9] Compared with traditional tissue biopsies, ctDNA can be obtained from blood samples, offering convenience and repeatability.[10,11] Studies have indicated that dynamic changes in ctDNA may correlate with the efficacy of *HER2*-targeted therapies; however, its role as a consistent predictive biomarker remains controversial, particularly with respect to different *HER2*-targeted agents.[12–14]

Amplification of *HER2* (also known as *ERBB2*) in breast cancer, as detected through tissue biopsy, has long been established as a hallmark of aggressive disease and poor prognosis. Similarly, previous studies have suggested that *PIK3CA* mutations identified in tissue biopsies may contribute to resistance to *HER2*-targeted therapies and adverse clinical outcomes.[15–18] However, the clinical implications of *HER2* amplification and *PIK3CA* mutations assessed through ctDNA remain uncertain. Unlike tissue-based findings, the predictive and prognostic value of ctDNA-detected genetic alterations, including *HER2* amplification and *PIK3CA* mutations, lacks robust validation. These uncertainties highlight the need for further evaluation of ctDNA mutations in the context of *HER2*-targeted therapies to better understand their clinical relevance.

Objectives

This meta-analysis aims to systematically evaluate the prognostic significance of ctDNA mutations in advanced *HER2*-positive breast cancer treated with *HER2*-targeted therapies. Specifically, we aim to explore how different gene mutations influence treatment outcomes and to clarify the predictive value of ctDNA mutations across diverse therapeutic strategies. Our findings may support clinical decision-making and contribute to the advancement of precision medicine in advanced *HER2*-positive breast cancer.

Materials and methods

Protocol and registration

This study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Two researchers (J.L. and H.X.) developed the protocol for this study, which was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (registration No. CRD42025638752).

Search strategy

Two researchers (J.L. and H.X.) systematically searched PubMed, Embase, and the Cochrane Library for relevant studies published before September 1, 2024. Additionally, abstracts from major annual meetings – including those of the American Society of Clinical Oncology (ASCO), the European Society for Medical Oncology (ESMO), and the San Antonio Breast Cancer Symposium (SABCS) – were manually reviewed for the period from March 2022

to September 2024. Search terms included Medical Subject Headings (MeSH) such as "breast cancer," "HER2," "TKI," "ADC," and "circulating tumor DNA," along with specific drug names and agents within these categories.

Selection criteria

The inclusion criteria for this study were as follows: 1) original research articles or conference abstracts presenting primary data, including observational studies (prospective or retrospective), randomized controlled trials (RCTs), or cross-sectional studies; 2) studies enrolling breast cancer patients treated with anti-*HER2* therapies, in which treatment regimens incorporated *HER2*-targeted agents such as trastuzumab, TKIs, or ADCs; 3) documented assessment of ctDNA, with no restrictions on detection or analytical methods due to the absence of a universally accepted standard; 4) reporting of ctDNA positivity rates along with relevant clinical outcomes, specifically progression-free survival (PFS) and/or overall survival (OS); and 5) publications available in English.

To ensure consistency across patient populations, studies were included only if *HER2*-positive breast cancer was defined according to ASCO/CAP guidelines, specifically as immunohistochemistry (IHC) 3+ or *HER2* gene amplification confirmed with in situ hybridization (ISH). Only studies involving administration of *HER2*-targeted agents, including monoclonal antibodies, TKIs, or ADCs, were considered eligible.

Exclusion criteria were as follows: 1) non-original publications, such as reviews, commentaries, or editorials; studies including fewer than 5 participants; or ongoing trials lacking publicly available results; 2) studies assessing only cell-free DNA (cfDNA) without specific analysis of ctDNA; and 3) investigations not involving *HER2*-targeted therapies.

Data extraction

Two researchers (J.L. and H.X.) independently extracted the following data from each eligible study: authors, publication year, study location, article format, study design, patient sample size, trial phase, details of the administered *HER2*-targeted regimen, and reported outcomes for patients with wild-type (WT) and mutant-type (MT) ctDNA receiving anti-*HER2* treatment. Any disagreements between the 2 researchers were resolved through consultation with a 3rd investigator.

Quality assessment and statistical analysis

Study quality was assessed using the Newcastle–Ottawa Scale (NOS), covering selection, comparability, and outcome domains. Scores of 6–9, 3–5, and 0–2 indicated low, moderate, and high risk of bias, respectively.[19]

Because clinical and methodological heterogeneity was anticipated among the included studies (differences in populations, treatment lines, ctDNA assays, and follow-up durations), a random-effects model (DerSimonian–Laird τ^2 estimator with Hartung–Knapp adjustment for confidence intervals) was pre-specified as the primary analysis model for all meta-analyses. Consistent with standard meta-analytic methodology, model selection was determined a priori rather than based on heterogeneity tests (I^2 or Q). For comparison and sensitivity analyses, fixed-effect models (Mantel–Haenszel method) were additionally calculated and reported when $I^2 < 25\%$ to illustrate result consistency under different assumptions.

Pooled hazard ratios (HRs) and odds ratios (ORs) with 95% confidence intervals (95% CIs) were calculated. When HRs were not directly reported, they were estimated from Kaplan–Meier curves using Tierney's method.[20–22] Between-study heterogeneity was quantified using Cochran's Q test ($p < 0.10$) and Higgins' I^2 statistic.

Sensitivity analyses were conducted for all meta-analyses, including the overall and subgroup analyses, by sequentially omitting individual studies (leave-one-out approach) to evaluate the robustness of the findings. Publication bias was assessed using funnel plots for all meta-analyses. Egger's regression test was performed when ≥10 studies were included; otherwise, only visual inspection was conducted.

All analyses were performed using R v. 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria) with the meta (v6.5-0), metafor (v4.2-0), and ggplot2 (v3.5.0) packages (https://CRAN.R-project.org/package=meta; https://CRAN.R-project.org/package=metafor; https://CRAN.R-project.org/package=ggplot2). Review Manager v. 5.3 (The Nordic Cochrane Centre, Cochrane Collaboration, Copenhagen, Denmark) was also used for supplementary analyses. Two-sided $p < 0.05$ were considered statistically significant, and an upper 95% CI < 1.0 indicated a favorable prognostic effect of WT ctDNA.

Results

Searching and screening of studies

A total of 560 records were identified through the preliminary systematic literature search, and 45 duplicate records were automatically identified and removed. After assessment and screening, 515 records were considered irrelevant and excluded. Among the 12 included reports, 11 were published full-text articles and one was an abstract presented at an annual conference, comprising a total of 558 patients.[23–34] The study selection process is depicted in Fig. 1. The flow diagram was developed in accordance with the PRISMA guidelines.[35]

The NOS appraisal is illustrated in Fig. 2 (study-level "risk-of-bias summary") and Fig. 3 (domain-level "risk-of-bias graph"). More than 80% of the judgments for core domains – such as representativeness of the exposed cohort, ascertainment of exposure, and outcome assessment – were rated

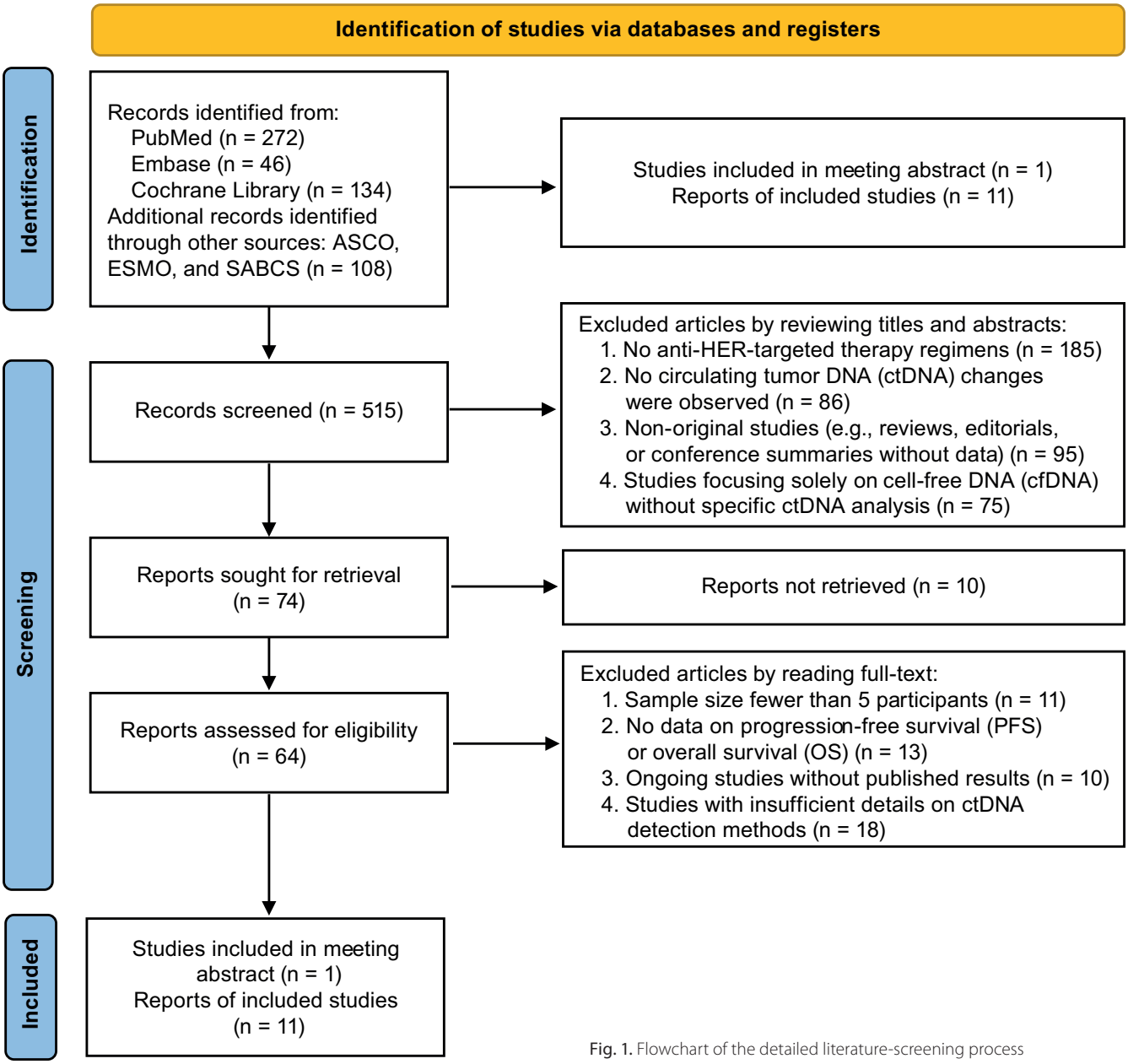


Fig. 1. Flowchart of the detailed literature-screening process

as low risk (green bars). By contrast, approx. 40% of the ratings for selection of the non-exposed cohort fell into the unclear-risk category (yellow bars), reflecting the frequent absence of an explicitly matched comparison group. Truly high-risk assessments were rare, accounting for no more than approx. 10% of the scores and occurring mainly in the item related to control for additional confounders. Overall, these data indicate that the methodological quality of the included studies was acceptable: most investigations satisfied the principal NOS criteria, although future studies should improve cohort comparability and confounder adjustment.

Characteristics of the included studies

Supplementary Table 1 presents the characteristics and outcomes of 12 included studies encompassing a total of 558 patients. Most studies were retrospective or cohort studies (n = 10), with a median sample size of 42 (range: 16–107). Treatment regimens varied widely and included trastuzumab (n = 1 study), TKIs (n = 8 studies), and ADCs (T-DM1; n = 3 studies), frequently combined with capecitabine or taxanes. Circulating tumor DNA mutations were detected in 8.57% to 84.20% of patients. Patients with ctDNA mutations consistently exhibited shorter PFS, with HRs ranging from 0.33 to 5.62 across different therapies. In contrast, WT ctDNA was associated with better clinical outcomes, highlighting its prognostic significance.

Prognostic association of ctDNA mutation status with PFS in *HER2*-targeted therapy

As shown in Fig. 4, the forest plot illustrates the association between ctDNA mutation status and PFS in patients receiving *HER2*-targeted therapy. Using a pre-specified

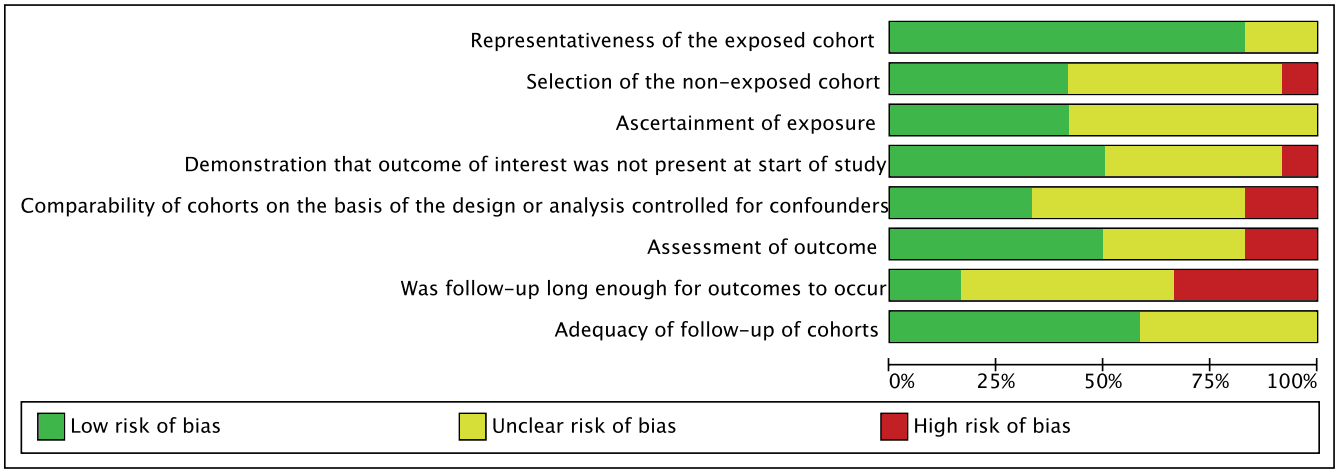


Fig. 2. Risk-of-bias summary for all included studies

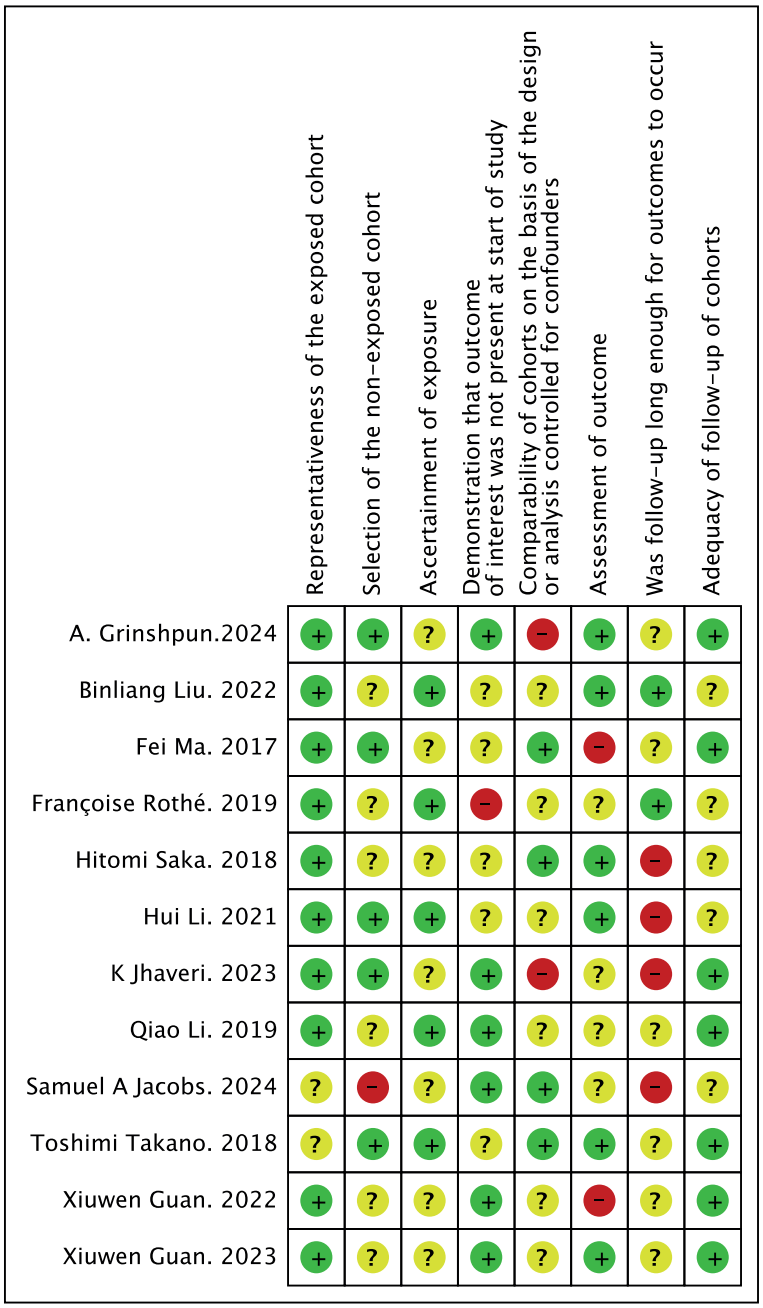


Fig. 3. Risk-of-bias graph for all included studies

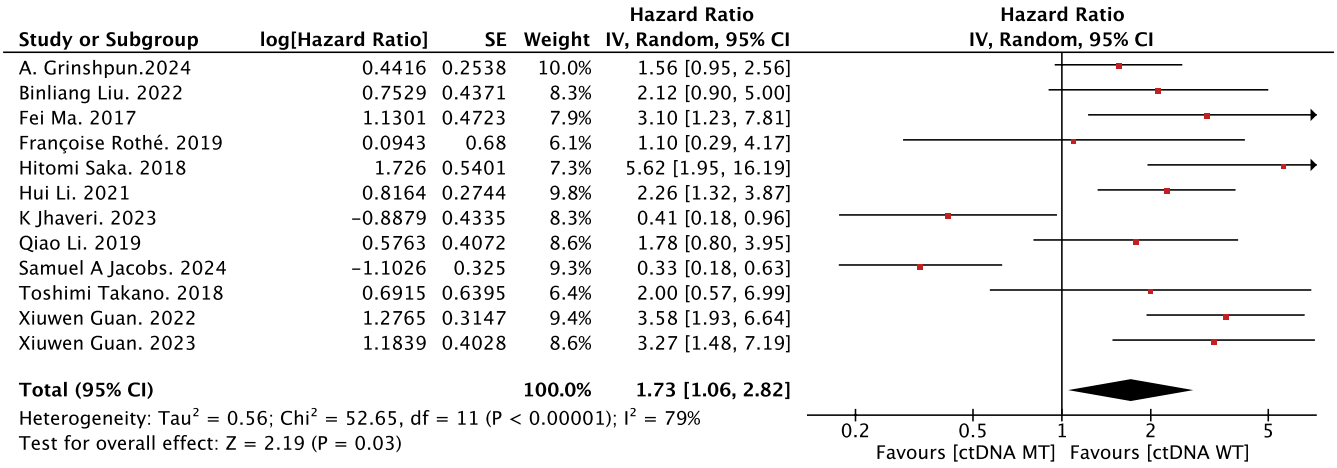


Fig. 4. Forest plot of progression-free survival (PFS) associated with circulating tumor DNA (ctDNA) mutations

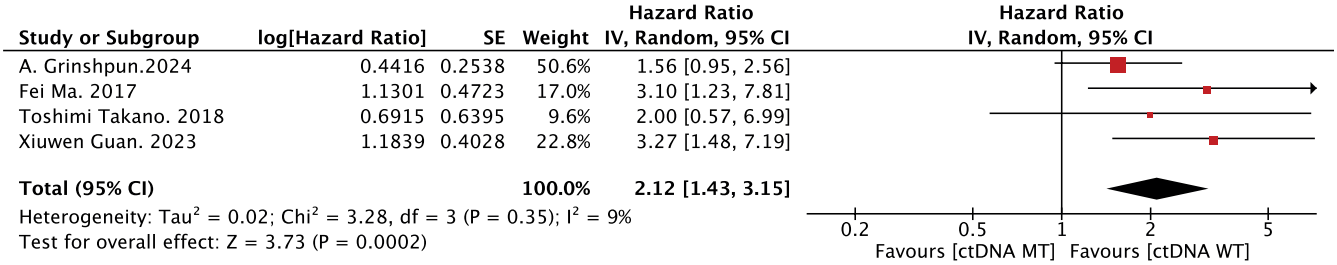


Fig. 5. Progression-free survival (PFS) in patients with PIK3CA mutation treated with anti-HER2 targeted therapy

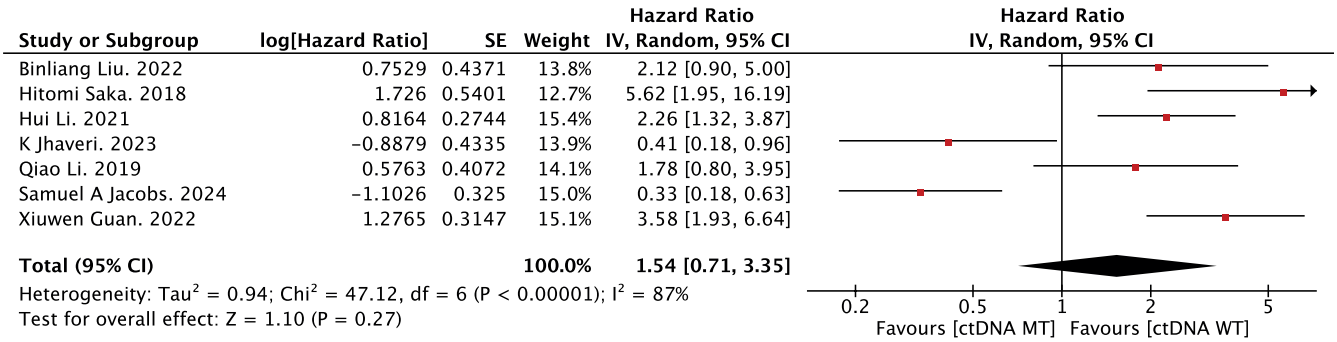


Fig. 6. Progression-free survival (PFS) in patients with *ERBB2* mutation treated with anti-HER2 targeted therapy

random-effects model, the pooled HR was 1.73 (95% CI: 1.06–2.82, p = 0.03), indicating that patients with ctDNA mutations had significantly shorter PFS compared with those harboring WT ctDNA.

Prognostic impact of ctDNA mutation subtypes in *HER2*-positive breast cancer

As shown in Fig. 5, *PIK3CA* mutations were significantly associated with reduced PFS in patients receiving *HER2*-targeted therapy (pooled HR = 2.12, 95% CI: 1.43–3.15, p = 0.002). This finding suggests that *PIK3CA* mutation status may serve as a biomarker for predicting early disease progression during *HER2*-targeted therapy. In contrast, *ERBB2* mutation status showed no statistically significant association with PFS (HR = 1.54, 95% CI: 0.71–3.35, p = 0.27; Fig. 6).

Prognostic value of ctDNA mutations in TKI-based treatment regimens

The prognostic role of ctDNA mutations was further analyzed in the context of TKI regimens (Fig. 7). In the overall TKI-treated cohort (n = 8 studies), ctDNA mutation positivity was significantly associated with shorter PFS (HR = 2.04, 95% CI: 1.29–3.24, p = 0.002), whereas no such association was observed in non-TKI-treated patients (HR = 1.29, 95% CI: 0.41–3.99, p = 0.66). Notably, this predictive effect differed substantially between TKI subtypes. For patients receiving pyrotinib

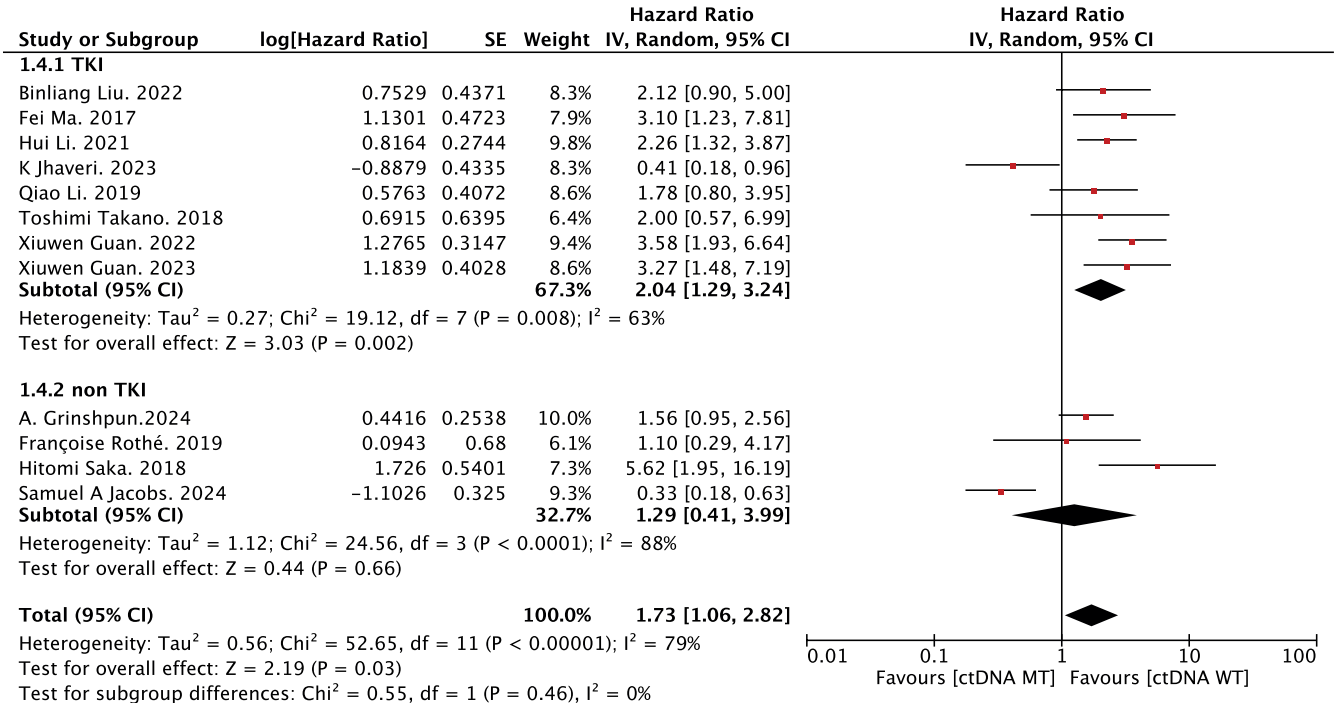


Fig. 7. Circulating tumor DNA (ctDNA) mutation status and progression-free survival (PFS) according to tyrosine kinase inhibitor (TKI) treatment status

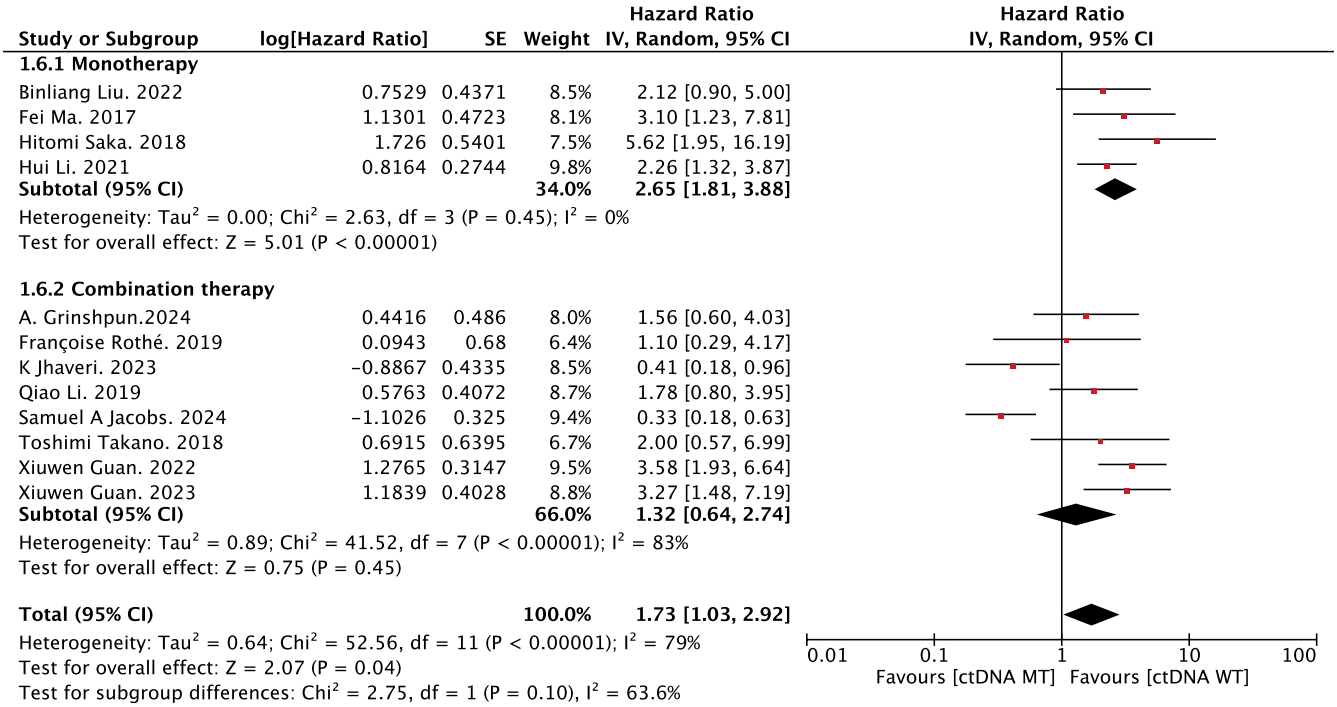


Fig. 8. Circulating tumor DNA (ctDNA) mutation status and progression-free survival (PFS) according to different tyrosine kinase inhibitor (TKI) regimens

(n = 5 studies), ctDNA mutations strongly predicted reduced PFS (HR = 2.77, 95% CI: 1.96–3.92, p < 0.001), with consistent directional trends across studies (Fig. 8). In contrast, non-pyrotinib TKIs (e.g., lapatinib and neratinib; n = 3 studies) showed no significant association between ctDNA status and PFS (HR = 1.23, 95% CI: 0.38–3.91, p = 0.73). These findings demonstrate subtype-specific associations between ctDNA mutations and PFS in TKI-treated cohorts, with the strongest effect observed in the pyrotinib subgroup.

Prognostic significance of ctDNA mutations across treatment combinations

The prognostic value of ctDNA mutations was evaluated across treatment strategies (Fig. 9). In monotherapy

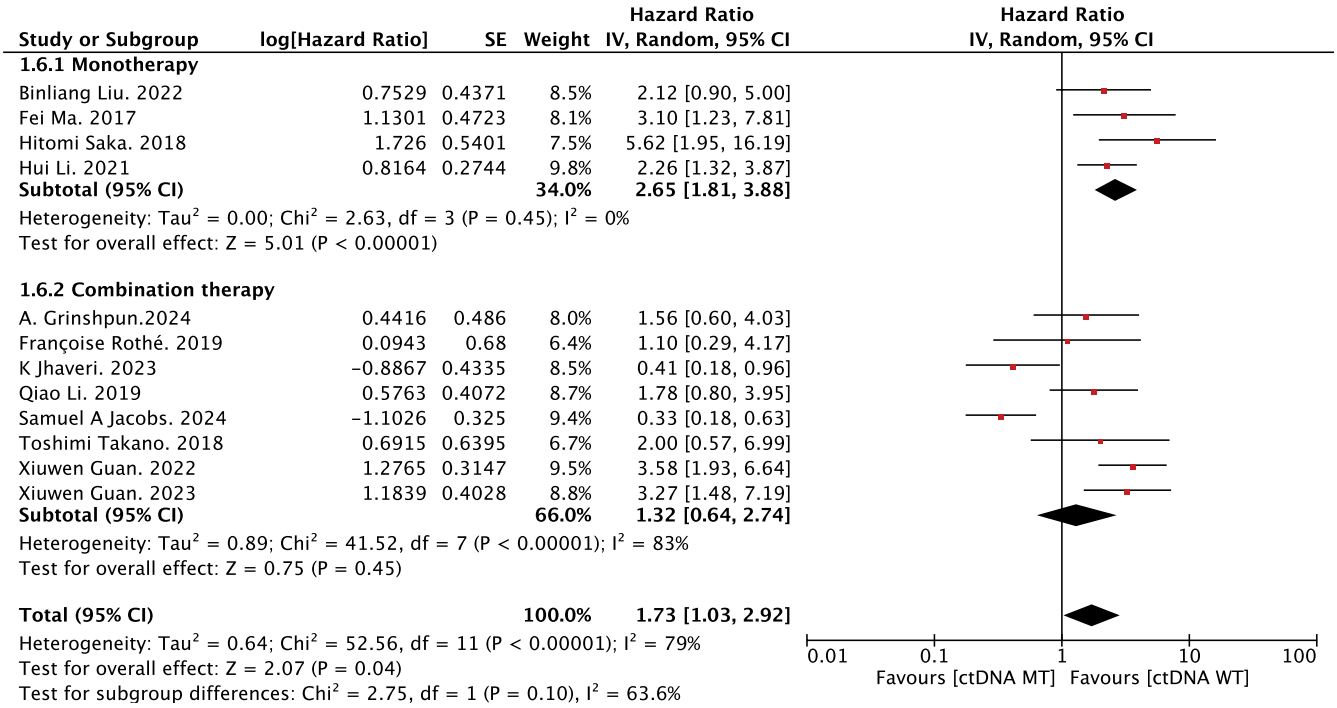


Fig. 9. Circulating tumor DNA (ctDNA) mutation status and progression-free survival (PFS) according to treatment strategy

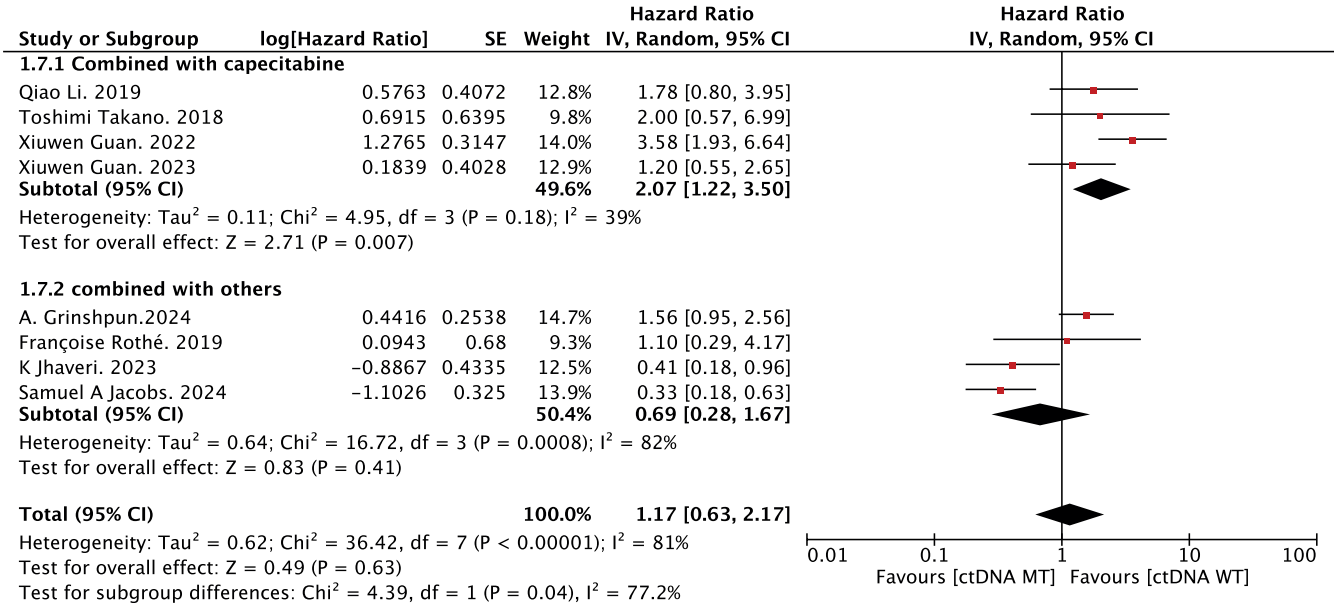


Fig. 10. Circulating tumor DNA (ctDNA) mutation status and progression-free survival (PFS) according to different combination regimens

cohorts, ctDNA mutation positivity was strongly associated with shorter PFS compared with WT cases (HR = 2.65, 95% CI: 1.81–3.88, p < 0.001). In contrast, combination therapy overall showed no significant difference in PFS between mutation-positive and mutation-negative groups (HR = 1.32, 95% CI: 0.64–2.74, p = 0.45). Further stratification of combination regimens revealed divergent effects. For capecitabine-based combinations (n = 4 studies), ctDNA mutations predicted markedly reduced PFS (HR = 2.07, 95% CI: 1.22–3.50, p = 0.007), whereas non-capecitabine combinations (e.g., taxane- or paclitaxel-based regimens; n = 4 studies) showed no association (HR = 0.69, 95% CI: 0.28–1.67, p = 0.41; Fig. 10).

Sensitivity and publication-bias analyses for the overall and subgroup meta-analyses

The leave-one-out sensitivity analysis demonstrated that sequential omission of any single study did not materially change the effect direction or statistical significance of the pooled estimate (Supplementary Fig. 1). The overall

estimates remained stable and within a narrow confidence interval (95% CI: 1.01–2.97), with no significant shifts in the hazard ratio (HR = 1.73). Publication bias assessment is shown in Supplementary Fig. 2. Egger's linear regression test detected no significant funnel plot asymmetry (t = 0.32, degrees of freedom (df) = 10, p = 0.796), indicating no evident small-study effects; the bias estimate was 0.7605 (standard error (SE) = 2.3438).

For all 10 predefined subgroups, the same leave-one-out sensitivity analyses and funnel plot evaluations were performed (Supplementary Fig. 3–22; mapping: *ERBB2* (Supplementary Fig. 3,4), *PIK3CA* (Supplementary Fig. 5,6), TKI (Supplementary Fig. 7,8), non-TKI (Supplementary Fig. 9,10), pyrotinib (Supplementary Fig. 11,12), non-pyrotinib (Supplementary Fig. 13,14), monotherapy (Supplementary Fig. 15,16), combination therapy (Supplementary Fig. 17,18), capecitabine-based combinations (Supplementary Fig. 19,20), and other combinations (Supplementary Fig. 21,22). Across subgroups, the pooled HRs remained directionally consistent, and significance judgments were not materially altered after omission of any single study, supporting the robustness of the findings. Because each subgroup included ≤10 studies, visual inspection of funnel plots was performed without formal Egger testing; no obvious asymmetry was observed.

Discussion

This meta-analysis evaluated the clinical significance of ctDNA mutations in assessing the efficacy of anti-*HER2* therapies in breast cancer, providing important insights into their potential role as biomarkers of therapeutic response and prognosis. The pooled analysis demonstrated a significant association between ctDNA mutations and shorter PFS across *HER2*-targeted therapies. These findings align with previous studies reporting that ctDNA mutations correlate with poor clinical outcomes, reinforcing their potential role in guiding personalized treatment strategies.[36] Among these, *PIK3CA* mutations showed a strong correlation with reduced PFS, which is consistent with existing evidence suggesting that activation of the PI3K/AKT/mTOR signaling pathway contributes to resistance to *HER2*-targeted therapies.[37]

This underscores the clinical relevance of detecting *PIK3CA* mutations via liquid biopsy as predictive biomarkers and suggests that patients harboring these mutations may benefit from treatment intensification, such as the incorporation of *PI3K* inhibitors. Notably, several ongoing clinical trials (e.g., NCT04208178 and NCT02038010) are investigating combinations of *HER2*-targeted agents and *PI3K* inhibitors, supporting the concept that dual blockade may help overcome resistance and improve patient outcomes. A key finding of this study is that the predictive value of ctDNA mutations was particularly evident in patients receiving TKI therapy, especially those treated with pyrotinib, suggesting that ctDNA analysis may assist in optimizing patient selection and informing timely treatment adjustments in this context. In contrast, this predictive effect was not observed in patients receiving T-DM1, which is consistent with previous studies reporting the limited predictive value of ctDNA mutations in this treatment setting.[38,39] This may be attributed to the cytotoxic mechanism of T-DM1, which depends predominantly on intracellular drug release rather than *HER2* signaling inhibition, thereby diminishing the impact of *PIK3CA* mutations on therapeutic efficacy.

Additionally, ctDNA mutations exhibited stronger predictive value in monotherapy than in combination therapies, where their predictive role was less pronounced. This may be attributed to the ability of combination treatments to counteract the effects of single-gene alterations through multiple mechanisms. However, in capecitabine-based regimens, ctDNA mutations remained a significant predictor of shorter PFS, suggesting that certain resistance mechanisms persist in this therapeutic setting.

Although *ERBB2* mutations detected in tissue biopsies have been well established as markers of resistance and poor prognosis in *HER2*-positive breast cancer, our analysis did not demonstrate a significant association between *ERBB2* mutations identified in ctDNA and PFS. This discrepancy likely reflects the inherent limitations of liquid biopsy compared with tissue-based methods. While tissue biopsy offers a static, localized snapshot of tumor genomic alterations, ctDNA reflects dynamic tumor burden and clonal evolution, which may be influenced by tumor heterogeneity, variable cfDNA shedding, assay sensitivity, and the type of therapy administered (such as TKIs or ADCs).[40,41] The use of pan-*HER* TKIs, particularly pyrotinib, may partially overcome downstream *HER2* pathway alterations, thereby mitigating the negative prognostic impact of *ERBB2* mutations detected in plasma. Additionally, methodological heterogeneity, variability in therapeutic regimens, and limited sample sizes in the included studies could have contributed to these findings. Future research involving larger cohorts, standardized ctDNA detection protocols, and stratification by treatment type will be necessary to clarify the predictive value of *ERBB2* mutations in liquid biopsy. Nevertheless, ctDNA testing offers unique advantages over tissue biopsy, including its noninvasive nature, the ability to capture real-time tumor dynamics, and the potential for longitudinal monitoring of resistance mechanisms in clinical practice.[42,43]

From a clinical perspective, liquid biopsy offers distinct advantages over traditional tissue biopsy by enabling real-time and repeatable assessment of treatment response and resistance development. This study supports the potential role of ctDNA in guiding treatment decisions, particularly in TKI- and pyrotinib-treated patients. However, the predictive value of ctDNA varies across treatment regimens; therefore, integration of multiple biomarkers may improve predictive accuracy in the future.[44,45] Additionally, previous

studies have reported that ctDNA detection rates correlate with metastatic tumor burden, with higher detection frequencies observed in patients with multiple or visceral metastases.[46–49] Our meta-analysis focused on the predictive role of ctDNA mutations and did not include tumor burden analysis. Future research should consider incorporating this factor to enhance the clinical utility of ctDNA. Future studies should also focus on standardizing ctDNA detection methods, validating these findings through prospective clinical trials, and leveraging longitudinal monitoring to refine precision oncology approaches for *HER2*-positive breast cancer.

Limitations of the study

Several limitations should be acknowledged. First, the included studies predominantly enrolled patients with advanced *HER2*-positive breast cancer, but variations in disease stage and prior treatments may have introduced heterogeneity. Second, ctDNA detection methods varied across studies, potentially affecting mutation rates and prognostic interpretation. Third, limited sample sizes and inconsistent reporting of baseline tumor burden reduced the ability to perform subgroup analyses for certain regimens. Fourth, *ERBB2* mutation analysis lacked statistical significance, possibly due to low mutation frequency and variable detection sensitivity. Finally, the observational nature of most included studies limits causal inference. Future prospective trials with standardized ctDNA methodologies and larger cohorts are needed to validate these findings in the context of advanced disease.

Conclusions

The mutational status of ctDNA holds promise as a prognostic biomarker in patients with *HER2*-positive breast cancer receiving targeted therapy. Specifically, *PIK3CA* mutations were consistently associated with worse survival outcomes, whereas *ERBB2* mutations did not demonstrate independent prognostic significance based on current evidence. The prognostic impact of ctDNA was most evident in patients treated with TKIs, particularly those receiving pyrotinib. Notably, this association remained significant in both monotherapy and capecitabine-based combination regimens. These findings support the use of ctDNA mutation profiling for risk stratification and personalized decision-making in the context of *HER2*-targeted therapy.

Supplementary data

The supplementary materials are available at: https://doi.org/10.5281/zenodo.17541159. The package contains the following files:

Supplementary Table 1. Summary of included studies evaluating the prognostic value of ctDNA mutation status in *HER2*-targeted therapy for breast cancer.

Supplementary Fig. 1. Leave-one-out sensitivity analysis for the overall PFS meta-analysis.

Supplementary Fig. 2. Funnel plot for publication bias in the overall PFS meta-analysis.

Supplementary Fig. 3. Sensitivity analysis of *ERBB2* subgroup.

Supplementary Fig. 4. Funnel plot of *ERBB2* subgroup.

Supplementary Fig. 5. Sensitivity analysis of *PIK3CA* subgroup.

Supplementary Fig. 6. Funnel plot of *PIK3CA* subgroup.

Supplementary Fig. 7. Sensitivity analysis of the TKI subgroup.

Supplementary Fig. 8. Funnel plot of the TKI subgroup.

Supplementary Fig. 9. Sensitivity analysis of the non-TKI subgroup.

Supplementary Fig. 10. Funnel plot of the non-TKI subgroup.

Supplementary Fig. 11. Sensitivity analysis of the pyrotinib subgroup.

Supplementary Fig. 12. Funnel plot of the pyrotinib subgroup.

Supplementary Fig. 13. Sensitivity analysis of the non-pyrotinib subgroup.

Supplementary Fig. 14. Funnel plot of the non-pyrotinib subgroup.

Supplementary Fig. 15. Sensitivity analysis of monotherapy subgroup.

Supplementary Fig. 16. Funnel plot of monotherapy subgroup.

Supplementary Fig. 17. Sensitivity analysis of combination subgroup.

Supplementary Fig. 18. Funnel plot of combination subgroup.

Supplementary Fig. 19. Sensitivity analysis of capecitabine-based subgroup.

Supplementary Fig. 20. Funnel plot of capecitabine-based subgroup.

Supplementary Fig. 21. Sensitivity analysis of other combination subgroup.

Supplementary Fig. 22. Funnel plot of other combination subgroup.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Jialin Lin https://orcid.org/0000-0002-6976-2172
Hangcheng Xu https://orcid.org/0000-0003-3576-6447
Jiayu Wang https://orcid.org/0000-0003-1967-389X
Binghe Xu https://orcid.org/0000-0002-0234-2747

References

1. Sobhani N, Roviello G, Corona SP, et al. The prognostic value of PI3K mutational status in breast cancer: A meta-analysis. *J Cell Biochem*. 2018;119(6):4287–4292. doi:10.1002/jcb.26687
2. De Blander H, Tonon L, Fauvet F, et al. Cooperative pro-tumorigenic adaptation to oncogenic RAS through epithelial-to-mesenchymal plasticity. *Sci Adv*. 2024;10(7):eadi1736. doi:10.1126/sciadv.adi1736
3. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*. 2010;12(5):R68. doi:10.1186/bcr2635
4. Pommier RM, Sanlaville A, Tonon L, et al. Comprehensive characterization of claudin-low breast tumors reflects the impact of the cell-of-origin on cancer evolution. *Nat Commun*. 2020;11(1):3431. doi:10.1038/s41467-020-17249-7
5. Veeraraghavan J, De Angelis C, Gutierrez C, et al. HER2-positive breast cancer treatment and resistance. *Adv Exp Med Biol*. 2025;1464:495–525. doi:10.1007/978-3-031-70875-6_24
6. Nami B, Ghanaeian A, Black C, Wang Z. Epigenetic silencing of HER2 expression during epithelial–mesenchymal transition leads to trastuzumab resistance in breast cancer. *Life*. 2021;11(9):868. doi:10.3390/life11090868
7. Gong B, Xue J, Yu J, et al. Cell-free DNA in blood is a potential diagnostic biomarker of breast cancer. *Oncol Lett*. 2012;3(4):897–900. doi:10.3892/ol.2012.576
8. Murtaza M, Dawson SJ, Tsui DWY, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature*. 2013;497(7447):108–112. doi:10.1038/nature12065
9. Dawson SJ, Tsui DWY, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med*. 2013;368(13):1199–1209. doi:10.1056/NEJMoa1213261
10. Fujita N, Nakayama T, Yamamoto N, et al. Methylated DNA and total DNA in serum detected by one-step methylation-specific PCR is predictive of poor prognosis for breast cancer patients. *Oncology*. 2012;83(5):273–282. doi:10.1159/000342083
11. Rui M, Wang Y, You JHS. Health economic evaluations of circulating tumor DNA testing for cancer screening: Systematic review. *Cancer Med*. 2025;14(3):e70641. doi:10.1002/cam4.70641
12. Garcia-Murillas I, Schiavon G, Weigelt B, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med*. 2015;7(302):302ra133. doi:10.1126/scitranslmed.aab0021
13. Olsson E, Winter C, George A, et al. Serial monitoring of circulating tumor DNA in patients with primary breast cancer for detection of occult metastatic disease. *EMBO Mol Med*. 2015;7(8):1034–1047. doi:10.15252/emmm.201404913
14. Park MS, Cho EH, Youn Y, et al. Importance of circulating tumor DNA analysis at diagnosis in early triple-negative breast cancer patients. *Breast Cancer*. 2025;32(2):416–425. doi:10.1007/s12282-025-01673-y
15. Loibl S, Majewski I, Guarneri V, et al. Correction to: *PIK3CA* mutations are associated with reduced pathological complete response rates in primary HER2-positive breast cancer: Pooled analysis of 967 patients from five prospective trials investigating lapatinib and trastuzumab. *Ann Oncol*. 2018;29(10):2151. doi:10.1093/annonc/mdx803
16. Fan H, Li C, Xiang Q, et al. *PIK3CA* mutations and their response to neoadjuvant treatment in early breast cancer: A systematic review and meta-analysis. *Thorac Cancer*. 2018;9(5):571–579. doi:10.1111/1759-7714.12618
17. Ibrahim EM, Kazkaz GA, Al-Mansour MM, Al-Foheidi ME. The predictive and prognostic role of phosphatase phosphoinositol-3 (PI3) kinase *(PIK3CA)* mutation in HER2-positive breast cancer receiving HER2-targeted therapy: A meta-analysis. *Breast Cancer Res Treat*. 2015;152(3):463–476. doi:10.1007/s10549-015-3480-6
18. Wang Y, Liu Y, Du Y, Yin W, Lu J. The predictive role of phosphatase and tensin homolog (PTEN) loss, phosphoinositol-3 (PI3) kinase (*PIK3CA*) mutation, and PI3K pathway activation in sensitivity to trastuzumab in HER2-positive breast cancer: A meta-analysis. *Curr Med Res Opin*. 2013;29(6):633–642. doi:10.1185/03007995.2013.794775
19. Stang A. Critical evaluation of the Newcastle–Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25(9):603–605. doi:10.1007/s10654-010-9491-z
20. Tierney JF, Burdett S, Fisher DJ. Practical methods for incorporating summary time-to-event data into meta-analysis: Updated guidance. *Syst Rev*. 2025;14(1):84. doi:10.1186/s13643-025-02752-z
21. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327(7414):557–560. doi:10.1136/bmj.327.7414.557
22. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statist Med*. 1998;17(24):2815–2834. doi:10.1002/(SICI)1097-0258(19981230)17:24<2815::AID-SIM110>3.0.CO;2-8
23. Jhaveri K, Eli LD, Wildiers H, et al. Neratinib + fulvestrant + trastuzumab for HR-positive, HER2-negative, HER2-mutant metastatic breast cancer: Outcomes and biomarker analysis from the SUMMIT trial. *Ann Oncol*. 2023;34(10):885–898. doi:10.1016/j.annonc.2023.08.003
24. Sakai H, Tsurutani J, Iwasa T, et al. HER2 genomic amplification in circulating tumor DNA and estrogen receptor positivity predict primary resistance to trastuzumab emtansine (T-DM1) in patients with HER2-positive metastatic breast cancer. *Breast Cancer*. 2018;25(5):605–613. doi:10.1007/s12282-018-0861-9
25. Takano T, Tsurutani J, Takahashi M, et al. A randomized phase II trial of trastuzumab plus capecitabine versus lapatinib plus capecitabine in patients with HER2-positive metastatic breast cancer previously treated with trastuzumab and taxanes: WJOG6110B/ELTOP. *Breast*. 2018;40:67–75. doi:10.1016/j.breast.2018.04.010
26. Rothé F, Silva MJ, Venet D, et al. Circulating tumor DNA in HER2-amplified breast cancer: A translational research substudy of the NeoALTTO phase III trial. *Clin Cancer Res*. 2019;25(12):3581–3588. doi:10.1158/1078-0432.CCR-18-2521
27. Guan X, Ma F, Li Q, et al. Pyrotinib monotherapy or pyrotinib in combination with capecitabine could significantly prolong progression-free survival and overall survival in patients with HER2-positive metastatic breast cancer. *J Clin Oncol*. 2022;40(16 Suppl):1034. doi:10.1200/JCO.2022.40.16_suppl.1034
28. Guan X, Ma F, Li Q, et al. Survival benefit and biomarker analysis of pyrotinib or pyrotinib plus capecitabine for patients with HER2-positive metastatic breast cancer: A pooled analysis of two phase I studies. *Biomark Res*. 2023;11(1):21. doi:10.1186/s40364-023-00453-0
29. Grinshpun A, Ren S, Graham N, et al. Phase Ib dose-escalation trial of taselisib (GDC-0032) in combination with HER2-directed therapies in patients with advanced HER2+ breast cancer. *ESMO Open*. 2024;9(6):103465. doi:10.1016/j.esmoop.2024.103465
30. Liu B, Yi Z, Guan Y, et al. Molecular landscape of *TP53* mutations in breast cancer and their utility for predicting the response to HER-targeted therapy in HER2 amplification-positive and HER2 mutation-positive amplification-negative patients. *Cancer Med*. 2022;11(14):2767–2778. doi:10.1002/cam4.4652
31. Ma F, Li Q, Chen S, et al. Phase I study and biomarker analysis of pyrotinib, a novel irreversible Pan-ErbB receptor tyrosine kinase inhibitor, in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer. *J Clin Oncol*. 2017;35(27):3105–3112. doi:10.1200/JCO.2016.69.6179
32. Li Q, Guan X, Chen S, et al. Safety, efficacy, and biomarker analysis of pyrotinib in combination with capecitabine in HER2-positive metastatic breast cancer patients: A phase I clinical trial. *Clin Cancer Res*. 2019;25(17):5212–5220. doi:10.1158/1078-0432.CCR-18-4173
33. Jacobs SA, Wang Y, Abraham J, et al. Correction: NSABP FB-10: A phase Ib/II trial evaluating ado-trastuzumab emtansine (T-DM1) with neratinib in women with metastatic HER2-positive breast cancer. *Breast Cancer Res*. 2024;26(1):83. doi:10.1186/s13058-024-01833-6
34. Li H, Wang J, Yi Z, et al. CDK12 inhibition enhances sensitivity of HER2+ breast cancers to HER2-tyrosine kinase inhibitor via suppressing PI3K/AKT. *Eur J Cancer*. 2021;145:92–108. doi:10.1016/j.ejca.2020.11.045
35. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Syst Rev*. 2022;18(2):e1230. doi:10.1002/cl2.1230
36. Jacob S, Davis AA, Gerratana L, et al. The use of serial circulating tumor DNA to detect resistance alterations in progressive metastatic breast cancer. *Clin Cancer Res*. 2021;27(5):1361–1370. doi:10.1158/1078-0432.CCR-20-1566

37. Fujimoto Y, Morita TY, Ohashi A, et al. Combination treatment with a PI3K/Akt/mTOR pathway inhibitor overcomes resistance to anti-HER2 therapy in *PIK3CA*-mutant HER2-positive breast cancer cells. *Sci Rep*. 2020;10(1):21762. doi:10.1038/s41598-020-78646-y
38. Baselga J, Cortés J, Im SA, et al. Biomarker analyses in CLEOPATRA: A phase III, placebo-controlled study of pertuzumab in human epidermal growth factor receptor 2-positive, first-line metastatic breast cancer. *J Clin Oncol*. 2014;32(33):3753–3761. doi:10.1200/JCO.2013.54.5384
39. Baselga J, Lewis Phillips GD, Verma S, et al. Relationship between tumor biomarkers and efficacy in EMILIA: A phase III study of trastuzumab emtansine in HER2-positive metastatic breast cancer. *Clin Cancer Res*. 2016;22(15):3755–3763. doi:10.1158/1078-0432.CCR-15-2499
40. Ding Y, Li W, Wang K, Xu C, Hao M, Ding L. Perspectives of the application of liquid biopsy in colorectal cancer. *Biomed Res Int*. 2020;2020(1):6843180. doi:10.1155/2020/6843180
41. Neumann MHD, Bender S, Krahn T, Schlange T. ctDNA and CTCs in liquid biopsy: Current status and where we need to progress. *Comput Struct Biotechnol J*. 2018;16:190–195. doi:10.1016/j.csbj.2018.05.002
42. Yin H, Zhang M, Zhang Y, Zhang X, Zhang X, Zhang B. Liquid biopsies in cancer. *Mol Biomed*. 2025;6(1):18. doi:10.1186/s43556-025-00257-8
43. Bronkhorst AJ, Ungerer V, Holdenrieder S. The emerging role of cell-free DNA as a molecular marker for cancer management. *Biomol Detect Quantif*. 2019;17:100087. doi:10.1016/j.bdq.2019.100087
44. Chang H, Anawate I, Low A, et al. Abstract PO4-14-10: Circulating tumor DNA as a biomarker for ADCs in metastatic breast cancer. *Cancer Res*. 2024;84(9 Suppl):PO4-14-10-PO4-14-10. doi:10.1158/1538-7445.SABCS23-PO4-14-10
45. Wang R, Wang B, Zhang H, et al. Early evaluation of circulating tumor DNA as marker of therapeutic efficacy and prognosis in breast cancer patients during primary systemic therapy. *Breast*. 2024;76:103738. doi:10.1016/j.breast.2024.103738
46. Raj R, Wehrle CJ, Aykun N, et al. Immunotherapy plus locoregional therapy leading to curative-intent hepatectomy in HCC: Proof of concept producing durable survival benefits detectable with liquid biopsy. *Cancers (Basel)*. 2023;15(21):5220. doi:10.3390/cancers15215220
47. Kahana-Edwin S, Torpy J, Cain LE, et al. Quantitative ctDNA detection in hepatoblastoma: Implications for precision medicine. *Cancers (Basel)*. 2023;16(1):12. doi:10.3390/cancers16010012
48. Wang Y, Wang W, Zhang T, et al. Dynamic bTMB combined with residual ctDNA improves survival prediction in locally advanced NSCLC patients with chemoradiotherapy and consolidation immunotherapy. *J Nat Cancer Center*. 2024;4(2):177–187. doi:10.1016/j.jncc.2024.01.008
49. Bharde A, Nadagouda S, Dongare M, et al. ctDNA-based liquid biopsy reveals wider mutational profile with therapy resistance and metastasis susceptibility signatures in early-stage breast cancer patients. *J Liq Biopsy*. 2025;7:100284. doi:10.1016/j.jlb.2024.100284

Retrospective analysis of guideline-based massage therapy in primary care for musculoskeletal disorders

Krzysztof Kassolik[1,A,B,F], Marcin Piwecki[2,C–E], Barbara Nowak[3,A,B,F], Ziemowit Nowak[3,B], Iwona Wilk[1,A,E], Jerzy Gielecki[4,A,F], Donata Kurpas[5,A,E,F]

[1] Faculty of Physiotherapy, Department of Fundamentals of Physiotherapy and Occupational Therapy, Wroclaw University of Health and Sport Sciences, Poland
[2] Doctoral School of Physical Culture Sciences, University of Physical Education in Kraków, Poland
[3] Barbara Nowak Family Doctor Practice, Wrocław, Poland
[4] Department of Anatomy and Histology, Faculty of Medicine, University of Warmia and Mazury in Olsztyn, Poland
[5] Faculty of Health Sciences, Wroclaw Medical University, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):969–977

Address for correspondence
Marcin Piwecki
E-mail: p665569523@gmail.com

Funding sources
None declared

Conflict of interest
None declared

Received on March 31, 2025
Reviewed on July 29, 2025
Accepted on August 29, 2025

Published online on June 10, 2026

Abstract

Background. Musculoskeletal disorders (MSDs) impose a significant burden on primary healthcare systems and the economy. Massage therapy (MT) may represent a useful tool for addressing these issues. However, the lack of robust evidence confirming its effectiveness makes this therapy controversial.

Objectives. This study aimed to assess the feasibility of implementing guideline-based MT into routine primary care practice.

Materials and methods. In this retrospective study, records of 258 primary care patients (median (Me) = 51.5 years; 1st and 3rd quartiles (Q1–Q3) = 42–66) were analyzed. These patients had previously received MT according to guidelines recommended by the Polish Society of Physiotherapy, the Polish Society of Family Medicine, the College of Family Physicians in Poland, and the European Rural and Isolated Practitioners Association (EURIPA). The effectiveness of therapy was evaluated using the visual analogue scale (VAS) and the number of general practitioner (GP) appointments booked by patients (NA).

Results. The majority of patients who received MT suffered from low back pain (M54.5) (39.1%); soft tissue disorders related to use, overuse, and pressure (M70) (27.5%); unspecified spondylosis (M47.9) (9.3%); and osteoarthritis of the knee (M17.5 and M17.9) (4.3%). A Wilcoxon test revealed a significant reduction in NA for M54.5 ($p < 0.001$), M70 ($p < 0.001$), others ($p < 0.001$), M17 ($p = 0.004$), and M47.9 ($p < 0.001$), as well as in VAS scores for M54.5 ($p < 0.001$), M70 ($p < 0.001$), others ($p < 0.001$), M47.9 ($p < 0.001$), and M17 ($p = 0.003$) after MT.

Conclusions. This study supports the potential benefits of integrating physiotherapy-led massage into routine primary care practice for the management of selected musculoskeletal disorders. When implemented in cooperation with family physicians, MT may reduce pain and decrease the need for additional GP visits.

Key words: primary health care, pain management, musculoskeletal diseases, clinical practice guidelines, massage therapy

Cite as
Kassolik K, Piwecki M, Nowak B, et al. Retrospective analysis of guideline-based massage therapy in primary care for musculoskeletal disorders. *Adv Clin Exp Med*. 2026;35(6):969–977. doi:10.17219/acem/210068

DOI
10.17219/acem/210068

Highlights

- Musculoskeletal disorders are a major burden for both patients and healthcare systems.
- Massage may be a valuable tool for the management of musculoskeletal disorders.
- The benefits of massage therapy remain understudied.
- Recommendations on massage approved by medical associations may facilitate its implementation in primary care settings.

Background

Musculoskeletal disorders (MSDs) are among the most prevalent and disabling health conditions worldwide, affecting approx. 1.71 billion people across all age groups.[1] These disorders not only cause persistent pain and limit mobility but also significantly impair quality of life and social functioning. Beyond their direct impact, MSDs are associated with an increased risk of developing chronic comorbidities such as cardiovascular disease (CVD), obesity, and diabetes.[2] The growing incidence of MSDs observed in recent decades reflects a combination of demographic changes, sedentary lifestyle, and occupational exposures.[3,4] This trend imposes an increasing burden on individuals, healthcare systems, and national economies.

Despite their prevalence, the management of MSDs remains suboptimal. Pharmacological interventions such as opioids and non-steroidal anti-inflammatory drugs (NSAIDs) are commonly used, particularly in early-stage disease.[5,6] However, the effectiveness of opioids in reducing musculoskeletal pain is questionable and may paradoxically contribute to prolonged work disability.[5] NSAIDs, although widely accepted, carry a well-documented risk of gastrointestinal, renal, and cardiovascular adverse effects.[6] This is especially concerning in the current healthcare landscape, where there is growing recognition of the public health consequences of long-term opioid and NSAID use.[7,8] As a result, there is an urgent need to identify safe and accessible non-pharmacological alternatives that could be implemented in primary healthcare settings.

In addition to their clinical consequences, MSDs generate substantial economic and systemic burdens. From the perspective of primary care, patients with MSDs frequently consult general practitioners (GPs), require diagnostic imaging, and often receive long-term pharmacotherapy.[9,10] These repeated interactions with the healthcare system result in increased costs, resource utilization, and administrative workload. Furthermore, MSDs contribute to absenteeism, reduced work productivity, and early retirement, particularly among working-age populations.[9] This underscores the need for cost-effective interventions that can reduce symptom burden while minimizing healthcare resource utilization.

Massage therapy (MT) is one such intervention. It is a well-established component of physiotherapeutic practice and has demonstrated promising results in relieving pain and improving function in patients with various MSDs, particularly in the early or subacute stages of disease progression.[11–13] Moreover, MT may be used as a preventive measure or as part of self-management strategies under professional guidance.[14–16] Despite its accessibility, safety, and low cost, MT remains underutilized in standard medical care.

There are 2 primary reasons for this underuse. First, the lack of consistent, high-quality empirical evidence supporting the clinical efficacy of MT limits its acceptance among medical professionals and hinders its integration into routine primary care protocols.[17–19] Second, the mechanisms through which MT exerts its therapeutic effects are not yet fully understood. Although evidence suggests that MT may influence biomechanical, neurological, and psychological processes, these effects remain insufficiently explained within a unified theoretical framework. Bridging this gap would require integration of disciplines such as anatomy, physiology, histology, and pathophysiology into MT research and training programs.[16,20] Importantly, the implementation of MT in primary care does not require highly specialized equipment or infrastructure. It may be delivered in a standard consultation room by a trained physiotherapist, either independently or in collaboration with family physicians. Such integration aligns with contemporary models of interdisciplinary care and is supported by national and international professional associations. Additionally, MT may serve as a form of autotherapy, contributing to self-management within the framework of social prescribing.[21] Nevertheless, real-world data regarding the practical implementation of MT in family medicine – particularly in terms of feasibility, effectiveness, and impact on healthcare utilization – remain limited.

This study addresses this gap by examining the use of guideline-based MT in patients with selected MSDs in a primary care setting. It focuses on the collaborative role of physiotherapists and GPs in delivering MT and its potential to reduce both pain and healthcare utilization.

Objectives

This study aimed to assess the feasibility of implementing guideline-based MT in routine primary care practice, in line with the recommendations of the Polish Society

of Physiotherapy, the Polish Society of Family Medicine, the College of Family Physicians in Poland, and the European Rural and Isolated Practitioners Association (EURIPA).

Specifically, the study explored whether MT, when delivered by physiotherapists in cooperation with family physicians, may contribute to reductions in both pain intensity and the number of GP consultations among patients with selected MSDs. The findings are intended to inform the potential integration of MT as a practical, non-pharmacological intervention into standard care pathways in family medicine.

Material and methods

Study design

This was a retrospective analysis of de-identified case records of primary care patients with selected MSDs. The project was conducted in accordance with the Declaration of Helsinki and was approved by the Senate Research Ethics Committee of the University of Physical Education in Wroclaw, Poland (approval No. 2/2018). The reporting of this study conforms to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.[22]

Setting

Data were collected between May 2019 and August 2023 at one of the medical centers in Wrocław. Data from patients treated with MT were obtained from the Dr Eryk medical database.

Participants

Participants' records were included in the dataset if they: 1) were aged 40 years or older; 2) were diagnosed with an MSD by a GP; 3) experienced pain greater than 3 on the visual analogue scale (VAS) before MT; 4) were prescribed MT interventions by a GP; and 5) had continuous medical records for at least 6 months before MT and no less than 6 months afterward at the time of data extraction from the database.

Patients' records were excluded from the dataset if they: 1) had a history of traumatic injuries, congenital musculoskeletal defects, deep vein thrombosis, myocardial infarction, or cancer within the past 5 years; 2) had undergone a surgical procedure within the past 5 years; 3) had acute inflammation of the respiratory, digestive, or genitourinary systems; 4) had incomplete medical records; or 5) were pregnant.

Variables

Pain assessment (primary endpoint)

Pain intensity was measured using the VAS. This is a continuous scale consisting of a 100-mm horizontal line anchored by "no pain" (score of 0) and "worst imaginable pain" (score of 10). The VAS is frequently used as an outcome measure to assess treatment effectiveness. Most studies have shown that the VAS is a valid and reliable scale.[23,24] Measurements were taken immediately before the intervention and after completion of the massage session.

Number of appointments (exploratory endpoint)

As an indicator of massage utility in primary healthcare, the number of appointments (NA) patients had with doctors during a 1-year period, including 6 months before and 6 months after MT, was analyzed. Consultations took place in doctors' offices. In every case, the analyzed change in NA referred to the same disorder. Patients who developed other diseases within 6 months after MT were excluded from the study. The reported reduction in NA was therefore directly associated with the original health condition under investigation and was not influenced by other health issues.[25]

Data sources

Part of the data (sex, age, type of disease, and NA) was extracted from the Dr Eryk database, which contains patient records entered by GPs from general practices across Poland. The patient population within the database may be considered representative of the Polish population in terms of demographics and disease distribution. The database contains patients' complete medical histories. In particular, information on the dates of registration with and withdrawal from the practice is collected. Information regarding referrals to secondary care, including specialty type, is also recorded. The database includes details on hospital admissions, discharge medications, diagnoses, outpatient consultations, investigations, and treatment outcomes. The database software is provided free of charge by the Polish National Health System. The authors did not have direct access to the database. Data on the number of massage sessions and pain intensity were extracted from the physiotherapy practice collaborating with the medical center operating the Dr Eryk database.

Bias

The NA indicator should not be considered a direct measure of massage efficacy. Rather, this indicator may suggest that MT is one of the factors positively influencing patients' health conditions. However, other factors that may affect patients' health status should also be taken into consideration. These include demographic factors (age, sex, education), psychosocial factors (beliefs, motivation), healthcare system factors (limited accessibility, long waiting times), socioeconomic factors (inability to take time off work,

cost, and income), disease-related factors (symptoms and disease severity), as well as tobacco smoking and alcohol consumption.[26]

Exposure

The MT intervention applied in this study followed a structured protocol developed and endorsed by the Polish Society of Physiotherapy, the Polish Society of Family Medicine, the College of Family Physicians in Poland, and EURIPA. The detailed protocol, consistently implemented by the physiotherapist, is publicly available.[16] To ensure adherence to the guidelines, a printed checklist was used during each session, listing the anatomical structures to be palpated and massaged. The therapist marked each completed step, thereby confirming adherence to the standardized procedure.

Unlike unstructured or practitioner-dependent massage practices, the protocol employed in this study was specifically designed for primary care settings. It required the therapist to target defined musculoskeletal structures – such as muscles, tendons, and ligaments – in a predetermined sequence. The approach aimed to restore functional balance in adjacent and interrelated anatomical areas. The protocol also included patient education and instruction in self-massage techniques to promote continuity of care and self-management between sessions. In contrast, unstructured MT often varies in technique, target areas, and therapeutic rationale depending on individual practitioner preferences or patient requests, which may limit reproducibility and consistency.

All treatments were delivered with the patient in a side-lying position, resting on the unaffected side. Supportive bolsters were placed under the head and upper and lower limbs to ensure comfort and proper alignment. The techniques used – effleurage, petrissage, friction, and compression – were based on Swedish massage methodology. According to Walton's massage therapy pressure scale, level 4 pressure (strong) was consistently applied.[27] Each session lasted 45 min. Massage was administered exclusively by a physiotherapist with more than 20 years of clinical experience and only when prescribed by a GP. Treatment continued until the patient reported a pain intensity of 3 or lower on the VAS, or until both the therapist and patient agreed to conclude therapy based on achievement of therapeutic goals and observed clinical progress.

Sample size

The sample size required to detect differences between 2 dependent means was calculated using G*Power software v. 3.1.9.7 (Heinrich Heine University (HHU), Düsseldorf, Germany). An effect size of Cohen's d = 0.25, a 2-sided $\alpha = 0.01$, and a statistical power of 90% were assumed. The required sample size was 242 participants. The number of extracted records was slightly higher than the predetermined sample size (n = 258).[28]

Statistical analyses

All statistical analyses were performed using IBM SPSS v. 29 (IBM Corp., Armonk, USA). Categorical variables are presented as frequencies and percentages (n, %). Continuous variables were assessed for normality of distribution using the Kolmogorov–Smirnov test (n > 50) or the Shapiro–Wilk test (n > 10) and are presented as mean ± standard deviation (SD) or median with interquartile range (Me; Q1–Q3), depending on the normality of distribution. Differences between measurements were analyzed using the χ^2 test or Fisher–Freeman–Halton exact test (FFH) when variables were categorical. Changes between 2 related samples were analyzed using the Wilcoxon signed-rank test when differences between measurements were not normally distributed, whereas a paired t-test was used when differences were normally distributed. For unrelated samples, the Mann–Whitney U test was used to compare 2 groups when variables were continuous and not normally distributed, while the Kruskal–Wallis H test was applied when more than 2 groups were compared. Effect sizes were calculated using r for the Mann–Whitney U test and Wilcoxon test ($r = Z/\sqrt{n}$), ϕ for the χ^2 test ($\phi = \sqrt{(\chi^2/n)}$), and Cramér's V for the FFH test. The significance level used for analysis was $p < 0.05$.

Results

Baseline characteristics of the participants are presented in Table 1. Briefly, 258 patients were reviewed (median age: 52 years; Q1–Q3 = 42–66 years); 159 (61.6%) were female and 99 (38.4%) were male. Based on the χ^2 goodness-of-fit test, the sex distribution in the sample did not match that of the population ($\chi^2(1) = 13.95$, $p < 0.001$, $\phi = 0.23$). A Mann–Whitney U test did not reveal a significant difference in age between women (Me = 54 years, n = 99) and men (Me = 49 years, n = 159) (U = 6644.5, p = 0.083, r = 0.108). On average, women were 5 years older than men. To analyze age as a potential variable, patients were divided into 3 age groups: 35 years and younger (15.5%), 36–55 years (39.5%), and 56 years and older (43.8%). A χ^2 goodness-of-fit test revealed a statistically significant difference between groups ($\chi^2(2) = 36.447$, $p < 0.001$, $\phi = 0.38$). Participants' body height was Me = 168 cm (Q1–Q3: 162–175 cm), body weight was Me = 76.5 kg (Q1–Q3: 65–87 kg), and body mass index (BMI) was Me = 24.77 kg/m^2 (Q1–Q3: 22.14–26.94 kg/m^2).

Most patients in the study exhibited chronic symptoms. The majority of patients who received MT suffered from M54.5 (low back pain; 39.1%), M70 (soft tissue disorders related to use, overuse, and pressure; 27.5%), M47.9 (spondylosis, unspecified; 9.3%), and osteoarthritis of the knee (M17.5 and M17.9; 4.3%) according to the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10).

Table 1. Baseline characteristics of participants

Variable		Women (n = 99)	Men (n = 159)	p-value
Age [years], Me (Q1–Q3)		54 (43–67)	49 (40–65)	0.083
Height [cm], Me (Q1–Q3)		168 (163–175)	168 (162–175)	0.712
Weight [kg], Me (Q1–Q3)		70.05 (61.41–77.81)	69.75 (59.65–78.77)	0.682
BMI [kg/m^2], Me (Q1–Q3)		24.93 (22.19–26.78)	24.55 (22.06–27.16)	0.794
ICD-10	M17.5 and M17.9, n (%)	8 (3.1)	3 (1.2)	0.212
	M47.9, n (%)	15 (5.8)	9 (3.5)	
	M54.5, n (%)	57 (22.1)	44 (17.1)	
	M70, n (%)	51 (19.8)	20 (7.8)	
	others, n (%)	28 (10.9)	23 (8.9)	

The χ^2 test was used for categorical data, and the Mann–Whitney U test was used for quantitative data; BMI – body mass index; Q1 and Q3 – 1st and 3rd quartile; Me – median; ICD-10 – International Statistical Classification of Diseases and Related Health Problems; International Classification of Diseases, 10th revision.

Table 2. Comparison of visual analogue scale (VAS) and number of appointments (NA) scores before and after massage therapy (MT) according to ICD-10 diagnostic categories

Diagnosis (ICD-10)	VAS before, Me (Q1–Q3)	VAS after, Me (Q1–Q3)	NA before, Me (Q1–Q3)	NA after, Me (Q1–Q3)
M17.5 and M17.9	5 (4–8)	1 (1–2)	1 (1–2)	0 (0–1)
M47.9	6 (4.5–8)	1 (1–2.5)	1 (1–2)	0 (0–1)
M54.5	7 (5–8)	1 (1–2)	1 (1–2)	0 (0–1)
M70	6 (5–8)	1 (1–2)	1 (1–2)	0 (0–1)
Others	7 (5–7)	1 (1–3)	1 (1–2)	0 (0–1)

ICD-10 – International Statistical Classification of Diseases and Related Health Problems; International Classification of Diseases, 10th revision.

The remaining cases were grouped into a category labeled "others" (19.8%). This group included the following ICD-10 codes: G54 (0.4%), M05 (0.4%), M06 (1.6%), M10 (0.8%), M15 (1.2%), M16 (1.9%), M18 (0.4%), M19 (1.2%), M41 (0.4%), M43 (0.8%), M48 (0.4%), M50 (2.3%), M53 (3.5%), M72 (0.4%), M75 (2.7%), M77 (0.8%), M79 (0.4%), and M81 (0.4%).

The results of the FFH test (p = 0.21, Cramér's V = 0.15) did not indicate a significant association between sex and disease type. Additionally, the FFH test (p = 0.025, Cramér's V = 0.189) demonstrated a significant association between disease type and age group.

The number of massage sessions

A total of 334 massage sessions were conducted in the analyzed group. The majority of patients (77.9%) experienced a significant reduction in pain after a single massage session (VAS < 3). Pain relief was achieved after 2 sessions in 17.1% of patients, while only 5.1% required 3 or more sessions for pain alleviation. Although the VAS threshold guided the intervention endpoint, therapy was limited to a maximum of 4 sessions to prevent indefinite continuation and ensure consistency across patients. A Mann–Whitney U test revealed the difference in the number of MT sessions between women (Me = 1, n = 99) and men (Me = 1, n = 159) in the group (U = 6824.5, p = 0.013, r = 0.154).

In terms of age, a Kruskal–Wallis H test revealed statistically significant differences between older patients who were 56 years old and older (Me = 1, n = 113), patients aged 36–55 years (Me = 1; n = 102) and patients aged 35 years and younger (Me = 1, n = 43) H (2) = 6.957, p = 0.031.

Visual analogue scale

The results of the VAS scores are presented in Table 2. Wilcoxon tests were used to compare VAS scores before and after MT. The results of the comparisons were reassessed using the Bonferroni correction. The analysis confirmed significant effects for all tested conditions: M54.5 (z = −8.664, p < 0.001, r = 0.862), M70 (z = −7.295, p < 0.001, r = 0.866), others (z = −6.124, p < 0.001, r = 0.857), M47.9 (z = −4.305, p < 0.001, r = 0.879), and M17 (z = −2.946, p = 0.003, r = 0.888) (Fig. 1).

The number of appointments

A total of 285 appointments were recorded within 6 months before MT. Every patient attended at least 1 appointment with a GP, with a maximum of 4 appointments. Following MT, a reduction in NA was observed (Table 2). Patients booked 127 appointments.

Patients were divided into 5 subgroups, and the results were compared using Wilcoxon tests. After applying the Bonferroni correction for multiple comparisons,

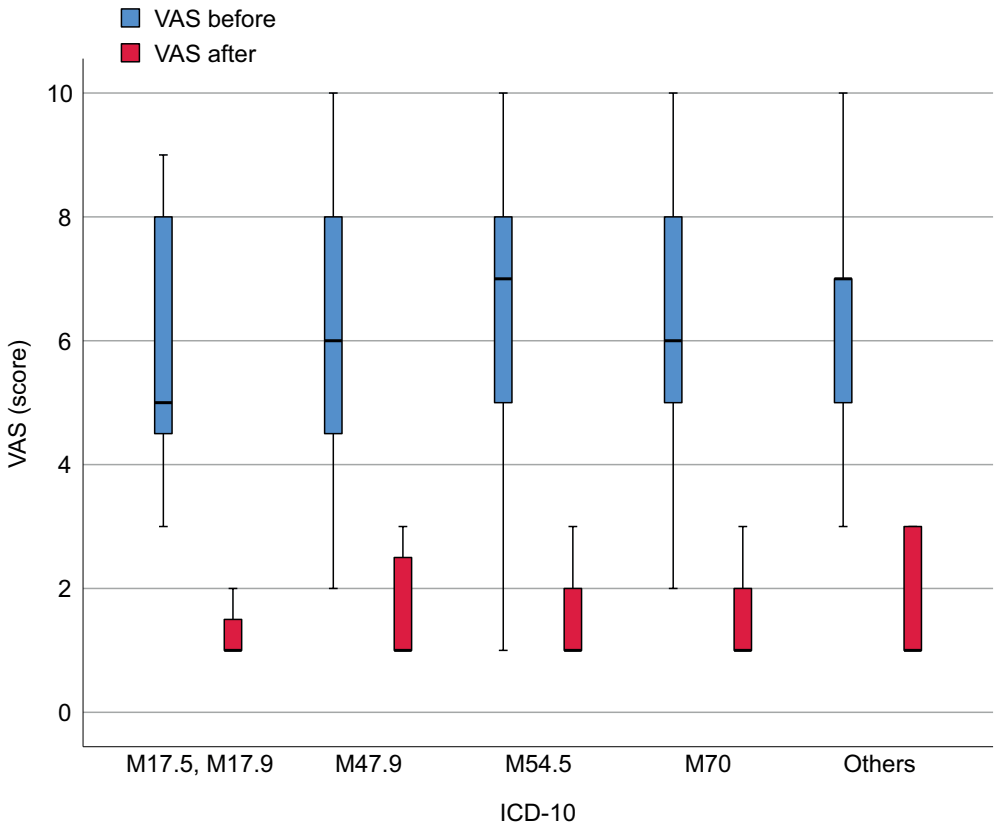


Fig. 1. Comparison of visual analogue scale (VAS) scores across different diseases

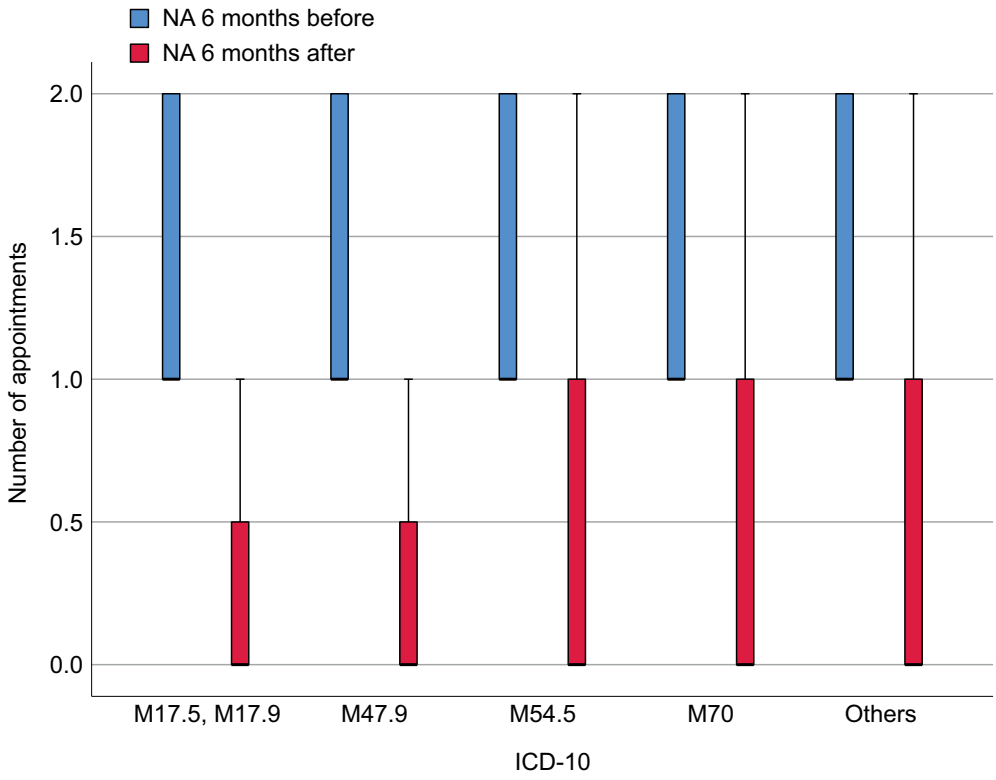


Fig. 2. Comparison of the number of appointments (NA) with a general practitioner (GP) across different diseases

the significance of the results was reassessed. M54.5 ($z = -7.978$, $p < 0.001$, $r = 0.794$), M70 ($z = -7.152$, $p < 0.001$, $r = 0.849$), others ($z = -5.124$, $p < 0.001$, $r = 0.717$), M17 ($z = -2.919$, $p = 0.004$, $r = 0.880$), and M47.9 ($z = -4.245$, $p < 0.001$, $r = 0.866$) remained highly significant (Fig. 2).

Discussion

This study assessed the implementation of guideline-based MT in a primary care setting among patients with selected MSDs, with particular focus on its association

with reductions in pain intensity and the number of GP appointments. The use of a structured protocol developed and endorsed by national and European professional societies, and applied in cooperation between physiotherapists and family physicians, makes this study a relevant contribution to the ongoing discussion regarding the integration of non-pharmacological therapies into routine primary care.

The key findings of this study – namely, the reduction in both pain intensity and GP visits – suggest a potential role for structured MT as a supportive intervention in the management of MSDs. These outcomes are particularly noteworthy given the limited number of comparable studies examining both clinical and organizational indicators. The observed effect should not be attributed solely to the physical intervention itself, but rather to the consistency and rigor of the applied protocol. This aligns with evidence from other therapeutic disciplines, where standardization and integration into broader care pathways tend to enhance treatment effectiveness.

Importantly, the findings must be interpreted in the context of mixed evidence regarding the efficacy of MT. Although some previous studies suggest that MT may lack long-term benefit or produce only modest effects,[29,30] these results often stem from trials using heterogeneous methods, unstructured treatment plans, or insufficient control of confounding factors. In contrast, the present study applied a uniform, anatomically informed protocol monitored through checklists and delivered in collaboration with GPs – conditions that more closely resemble the intended setting of real-world implementation.

Still, MT should not be seen as a universal or standalone solution. While a substantial proportion of patients benefited from therapy, some reported persistent pain requiring additional or alternative forms of management. In such cases, MT may serve best as an adjunct to pharmacotherapy, structured exercise, or psychological interventions. Conversely, in patients who cannot tolerate pharmacological options or who present with contraindications to standard treatments, MT may represent a valuable therapeutic alternative. This interpretation corresponds with World Health Organization (WHO) guidelines, which recommend massage only as part of a broader management plan for chronic low back pain,[31] and is also consistent with evidence of short- to medium-term efficacy reported by Furlan et al.[32]

The evidence supporting the effectiveness of MT in different MSD subtypes remains heterogeneous. For example, in knee osteoarthritis, Wu et al. and Perlman et al. provided differing assessments of MT, ranging from limited short-term benefit[33] to functional improvement and pain reduction.[34] Perlman et al. noted a clinically significant benefit using a whole-body massage protocol resembling the structured intervention applied in the present study.[34] Similarly, Jurecka et al. demonstrated that soft tissue therapy, including MT, may be useful both as a primary and adjunctive approach.[35] On the other hand, Yuan et al. concluded that the current evidence base remains inconclusive and that more rigorous research is needed.[36]

In spondylosis-related pain, findings are also variable. Cheng and Huang found moderate-quality evidence supporting the efficacy of MT in cervical spondylosis,[37] while Al-Mutairi et al. and Gross et al. raised concerns regarding the insufficient quality of available studies.[38,39] For thoracic spondylosis, data remain scarce, although Srokowska et al. demonstrated promising effects of deep tissue massage in the management of thoracic spine pain.[40] However, methodological differences between their protocol and the one used in the present study limit the comparability of the results.

The impact of MT on soft tissue disorders related to overuse, pressure, or poor posture has been relatively underexplored. While much of the literature focuses on athletes, primary care populations with work-related or degenerative soft tissue conditions remain understudied. Our findings suggest that MT may provide relief in these cases, particularly when pain is localized and not attributable to systemic inflammatory processes. Further research is warranted to validate its role in this subpopulation.[14]

From a broader perspective, the discussion regarding the utility of MT in primary care should also consider its feasibility and acceptability in routine practice. The protocol used in this study required minimal equipment and could be implemented in a standard medical consultation room. It was conducted by a physiotherapist with more than 20 years of experience and always upon GP referral. Although this enhances the internal validity of the findings, it may also limit generalizability, as not all practices have access to such highly trained personnel. Moreover, frequent application, as required in some chronic pain conditions, could place strain on healthcare resources and increase waiting times. As noted in previous studies, this logistical barrier may partially explain the limited effectiveness of MT in real-world settings.[29,30]

Finally, although this study focused primarily on clinical and healthcare utilization outcomes, it is important to acknowledge the broader psychosocial benefits of MT, including reductions in anxiety, increased satisfaction with care, and perceived support from healthcare providers.[41] These factors, although not formally assessed in the present analysis, may mediate or enhance the therapeutic effects of MT and merit inclusion in future studies.

This study provides preliminary yet promising evidence that structured, guideline-informed MT may be a feasible and potentially effective intervention for selected MSDs in primary care. When embedded in collaborative models involving GPs and physiotherapists, and when targeted at appropriate clinical indications, MT may offer both symptom relief and reductions in healthcare utilization. Further controlled studies with long-term follow-up and broader implementation contexts are needed to confirm these findings and inform policy and clinical guidelines.

Interpretation

The results suggest that MT, when provided by physiotherapists in collaboration with family physicians, may help reduce both pain intensity and the number of GP consultations in patients with selected musculoskeletal disorders. This highlights the potential value of including physiotherapeutic interventions directly in primary care settings. Although the retrospective nature of the study limits causal interpretation, the observed outcomes support further exploration of structured cooperation models between physiotherapists and primary care providers.

Generalizability

The study involved adult patients with chronic musculoskeletal complaints treated in a primary care setting. Massage therapy was delivered by a qualified physiotherapist operating in close coordination with a family physician, which reflects a feasible model of interdisciplinary care. However, the generalizability of the findings may be limited in contexts where such collaboration is not structurally supported or where physiotherapists are not integrated into primary care teams. The conclusions also pertain to a specific type of massage and cannot be extrapolated to all forms of manual therapy.

Limitations of the study

This study has several important limitations that should be considered when interpreting the findings. First, its retrospective design precludes causal inferences. As data were collected after the intervention had already taken place, the study is subject to potential selection and information biases, which may have influenced the observed associations. Second, the presence of uncontrolled confounding variables limits interpretation of the results. Patients may have used analgesics, participated in physical therapy, performed self-directed exercises, or received psychological support during the observation period. These co-interventions were not documented and could have contributed to the observed improvements independently of MT. Third, the number of GP NAs, used as an exploratory indicator of potential benefit, may be influenced by various non-clinical factors such as healthcare access, patient preferences, or organizational policies. Therefore, this metric should not be regarded as a direct measure of clinical effectiveness. Moreover, the study did not distinguish between acute and chronic presentations of musculoskeletal disorders, nor did it account for possible delayed effects of treatments initiated prior to the massage intervention. Finally, the generalizability of the findings is limited by the use of a single therapist following a specific protocol, and comparisons with other studies are constrained by methodological heterogeneity, including variations in massage techniques, diagnostic classifications, and outcome measures.

Conclusions

This study supports the potential integration of physiotherapy-led massage into routine primary care practice for the management of selected musculoskeletal disorders. When implemented in cooperation with family physicians, MT may reduce pain and decrease the need for additional GP visits. This underscores the importance of closer collaboration between physiotherapists and primary care providers in developing efficient, evidence-informed care pathways for patients with musculoskeletal complaints.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Figshare at https://doi.org/10.6084/m9.figshare.28716092.

Consent for publication

Not applicable.

Use of AI and AI-assisted technology

Not applicable.

ORCID iDs

Krzysztof Kassolik https://orcid.org/0000-0003-2836-3703
Marcin Piwecki https://orcid.org/0000-0002-6799-3997
Barbara Nowak https://orcid.org/0000-0001-6683-6351
Ziemowit Nowak https://orcid.org/0000-0003-3057-5545
Iwona Wilk https://orcid.org/0000-0003-4914-8391
Jerzy Gielecki https://orcid.org/0000-0002-3039-873X
Donata Kurpas https://orcid.org/0000-0002-6996-8920

References

1. Williams A, Kamper SJ, Wiggers JH, et al. Musculoskeletal conditions may increase the risk of chronic disease: A systematic review and meta-analysis of cohort studies. *BMC Med*. 2018;16(1):167. doi:10.1186/s12916-018-1151-2
2. Yang MH, Jhan CJ, Hsieh PC, Kao CC. A study on the correlations between musculoskeletal disorders and work-related psychosocial factors among nursing aides in long-term care facilities. *Int J Environ Res Public Health*. 2021;19(1):255. doi:10.3390/ijerph19010255
3. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10267):2006–2017. doi:10.1016/S0140-6736(20)32340-0
4. Lewis R, Gómez Álvarez CB, Rayman M, Lanham-New S, Woolf A, Mobasheri A. Strategies for optimising musculoskeletal health in the 21st century. *BMC Musculoskelet Disord*. 2019;20(1):164. doi:10.1186/s12891-019-2510-7
5. Carnide N, Hogg-Johnson S, Côté P, et al. Early prescription opioid use for musculoskeletal disorders and work outcomes: A systematic review of the literature. *Clin J Pain*. 2017;33(7):647–658. doi:10.1097/AJP.0000000000000452
6. Bindu S, Mazumder S, Bandyopadhyay U. Non-steroidal anti-inflammatory drugs (NSAIDs) and organ damage: A current perspective. *Biochem Pharmacol*. 2020;180:114147. doi:10.1016/j.bcp.2020.114147
7. Brent J, Weiss ST. The opioid crisis: Not just opioids anymore. *JAMA Netw Open*. 2022;5(6):e2215432. doi:10.1001/jamanetworkopen.2022.15432

8. Gardner EA, McGrath SA, Dowling D, Bai D. The opioid crisis: Prevalence and markets of opioids. *Forensic Sci Rev*. 2022;34(1):43–70. PMID:35105535.
9. Power JD, Perruccio AV, Paterson JM, et al. Healthcare utilization and costs for musculoskeletal disorders in Ontario, Canada. *J Rheumatol*. 2022;49(7):740–747. doi:10.3899/jrheum.210938
10. Bornhöft L, Thorn J, Svensson M, Nordeman L, Eggertsen R, Larsson MEH. More cost-effective management of patients with musculoskeletal disorders in primary care after direct triaging to physiotherapists for initial assessment compared to initial general practitioner assessment. *BMC Musculoskelet Disord*. 2019;20(1):186. doi:10.1186/s12891-019-2553-9
11. Chen PC, Wei L, Huang CY, Chang FH, Lin YN. The effect of massage force on relieving nonspecific low back pain: A randomized controlled trial. *Int J Environ Res Public Health*. 2022;19(20):13191. doi:10.3390/ijerph192013191
12. Kong LJ, Zhan HS, Cheng YW, Yuan WA, Chen B, Fang M. Massage therapy for neck and shoulder pain: A systematic review and meta-analysis. *Evid Based Complement Alternat Med*. 2013;2013:613279. doi:10.1155/2013/613279
13. Field T. Massage therapy research review. *Complement Ther Clin Pract*. 2014;20(4):224–229. doi:10.1016/j.ctcp.2014.07.002
14. Weerapong P, Hume PA, Kolt GS. The mechanisms of massage and effects on performance, muscle recovery and injury prevention. *Sports Med*. 2005;35(3):235–256. doi:10.2165/00007256-200535030-00004
15. Eghbali M, Lellahgani H, Alimohammadi N, Daryabeigi R, Ghasempour Z. Study on effect of massage therapy on pain severity in orthopedic patients. *Iran J Nurs Midwifery Res*. 2010;15(1):32–36. PMID:21589747. PMCID:PMC3093031.
16. Kassolik K, Rajkowska-Labon E, Wilk I, et al. Recommendations of the Polish Society of Physiotherapy, the Polish Society of Family Medicine, the College of Family Physicians in Poland and the European Rural and Isolated Practitioners Association regarding the use of simple forms of physiotherapy, including massage and self-massage in primary care, endorsed by the Polish Society of Physiotherapy Specialists. *Fam Med Prim Care Rev*. 2019;21(3):277–288. https://www.termedia.pl/Recommendations-of-the-Polish-Society-of-Physiotherapy-the-Polish-Society-of-Family-Medicine-the-College-of-Family-Physicians-in-Poland-and-the-European-Rural-and-Isolated-Practitioners-Association-re,95,37870,0,1.html.
17. Thomson D, Gupta A, Arundell J, Crosbie J. Deep soft-tissue massage applied to healthy calf muscle has no effect on passive mechanical properties: A randomized, single-blind, cross-over study. *BMC Sports Sci Med Rehabil*. 2015;7(1):21. doi:10.1186/s13102-015-0015-8
18. Miake-Lye IM, Mak S, Lee J, et al. Massage for pain: An evidence map. *J Altern Complement Med*. 2019;25(5):475–502. doi:10.1089/acm.2018.0282
19. Davis HL, Alabed S, Chico TJA. Effect of sports massage on performance and recovery: A systematic review and meta-analysis. *BMJ Open Sport Exerc Med*. 2020;6(1):e000614. doi:10.1136/bmjsem-2019-000614
20. Gerivani A, Sadeghi T, Karimi Moonaghi H, Zendedel A. Integrating of Anatomy and Physiology courses in basic medical sciences (case study in Mashhad Faculty of Medicine). *Future Med Educ J*. 2020;10(4):46–50. doi:10.22038/fmej.2020.48087.1327
21. Petrazzuoli F, Vidal-Alaball J, Kenkre J, et al. Best practice approaches to social prescribing in European Primary Care: A Delphi protocol focused on link workers. *Adv Clin Exp Med*. 2025;34(9):1589–1595. doi:10.17219/acem/208216
22. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806–808. doi:10.1136/bmj.39335.541782.AD
23. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken)*. 2011;63(Suppl 11):S240–S252. doi:10.1002/acr.20543
24. Begum R, Hossain MR. Validity and reliability of visual analogue scale (VAS) for pain measurement. *J Med Case Rep Rev*. 2019;2:11. https://jmcrr.info/index.php/jmcrr/article/view/44.
25. Organisation for Economic Co-operation and Development (OECD). Doctors' consultations. 2024. doi:10.1787/173dcf26-en
26. Jin J, Sklar GE, Min Sen Oh V, Chuen Li S. Factors affecting therapeutic compliance: A review from the patient's perspective. *Ther Clin Risk Manag*. 2008;4(1):269–286. doi:10.2147/tcrm.s1458
27. Walton T. *Medical Conditions and Massage Therapy: A Decision Tree Approach*. Philadelphia, USA: Wolters Kluwer/Lippincott Williams & Wilkins Health; 2011. ISBN:978-0-7817-6922-8.
28. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175–191. doi:10.3758/BF03193146
29. Bervoets DC, Luijsterburg PA, Alessie JJ, Buijs MJ, Verhagen AP. Massage therapy has short-term benefits for people with common musculoskeletal disorders compared to no treatment: A systematic review. *J Physiother*. 2015;61(3):106–116. doi:10.1016/j.jphys.2015.05.018
30. Kang T, Kim B. Cervical and scapula-focused resistance exercise program versus trapezius massage in patients with chronic neck pain: A randomized controlled trial. *Medicine (Baltimore)*. 2022;101(39):e30887. doi:10.1097/MD.0000000000030887
31. World Health Organization (WHO): Ageing and Health (AAH), Guidelines Review Committee, Maternal, Newborn, Child & Adolescent Health & Ageing (MCA). *WHO Guideline for Non-Surgical Management of Chronic Primary Low Back Pain in Adults in Primary and Community Care Settings*. Geneva, Switzerland: World Health Organization (WHO); 2023. ISBN:978-92-4-008178-9.
32. Furlan AD, Giraldo M, Baskwill A, Irvin E, Imamura M. Massage for low-back pain. *Cochrane Database Syst Rev*. 2015;2015(9):CD001929. doi:10.1002/14651858.CD001929.pub3
33. Wu Q, Zhao J, Guo W. Efficacy of massage therapy in improving outcomes in knee osteoarthritis: A systematic review and meta-analysis. *Complement Ther Clin Pract*. 2022;46:101522. doi:10.1016/j.ctcp.2021.101522
34. Perlman A, Fogerite SG, Glass O, et al. Efficacy and safety of massage for osteoarthritis of the knee: A randomized clinical trial. *J Gen Intern Med*. 2019;34(3):379–386. doi:10.1007/s11606-018-4763-5
35. Jurecka A, Papież M, Skucińska P, Gądek A. Evaluating the effectiveness of soft tissue therapy in the treatment of disorders and postoperative conditions of the knee joint: A systematic review. *J Clin Med*. 2021;10(24):5944. doi:10.3390/jcm10245944
36. Ma X, Yuan Z, Qian B, Guan Y, Wang B. Systematic review and meta-analysis of reflexology for people with multiple sclerosis. *Medicine (Baltimore)*. 2023;102(5):e32661. doi:10.1097/MD.0000000000032661
37. Cheng YH, Huang GC. Efficacy of massage therapy on pain and dysfunction in patients with neck pain: A systematic review and meta-analysis. *Evid Based Complement Alternat Med*. 2014;2014(1):204360. doi:10.1155/2014/204360
38. Al-Mutairi AMF, Almutairi STF, Alruaidan AR, Almutairi DM, Alqthami AMG. Cervical spondylosis: Innovative techniques for physical therapy: An updated review. *J Ecohumanism*. 2024;3(8):13195. doi:10.62754/joe.v3i8.6215
39. Gross A, Langevin P, Burnie SJ, et al. Manipulation and mobilisation for neck pain contrasted against an inactive control or another active treatment. *Cochrane Database Syst Rev*. 2015;2015(9):CD004249. doi:10.1002/14651858.CD004249.pub4
40. Srokowska A, Bodek M, Kurczewski M, Srokowski G, Lewandowski A. Deep tissue massage and mobility and pain in the thoracic spine. *Baltic J Health Phys Act*. 2019;11(2):99–108. doi:10.29359/BJHPA.11.2.10
41. Miller J, Dunion A, Dunn N, et al. Effect of a brief massage on pain, anxiety, and satisfaction with pain management in postoperative orthopaedic patients. *Orthop Nurs*. 2015;34(4):227–234. doi:10.1097/NOR.0000000000000163

Impact of primary tumor SUV_{max} on PET accuracy in mediastinal lymph node staging of non-small cell lung cancer (NSCLC)

Adam Kuźdżał[1,A–F], Karolina Gambuś[2,C,E,F], Konrad Moszczyński[3,B,E,F], Sofiia Popovchenko[4,B,E,F], Monika Bryndza[4,B,E,F], Lucyna Rudnicka[5,B,E,F], Janusz Warmus[6,B,E,F], Jolanta Hauer[6,B,E,F], Katarzyna Żanowska[6,B,E,F], Łukasz Trybalski[6,B,E,F], Piotr Kocoń[7,A–F]

[1] Department of Radiology, Maria Skłodowska-Curie National Institute of Oncology, National Research Institute, Cracow, Poland
[2] Department of Ophthalmology, Ludwik Rydygier Hospital, Cracow, Poland
[3] Department of Coronary Artery Disease and Structural Heart Disease, Institute of Cardiology, Warsaw, Poland
[4] Students' Scientific Society, Jagiellonian University Medical College, Cracow, Poland
[5] Department of Pathology, John Paul II Hospital, Cracow, Poland
[6] Department of Thoracic Surgery, John Paul II Hospital, Cracow, Poland
[7] Department of Thoracic Surgery, Jagiellonian University Medical College, John Paul II Hospital, Cracow, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):979–984

Address for correspondence
Adam Kuźdżał
E-mail: adam.kuzdzal@krakow.nio.gov.pl

Funding sources
None declared

Conflict of interest
None declared

Received on January 13, 2025
Reviewed on July 3, 2025
Accepted on September 5, 2025

Published online on June 25, 2026

Cite as
Kuźdżał A, Gambuś K, Moszczyński K, et al. Impact of primary tumor SUV_{max} on PET accuracy in mediastinal lymph node staging of non-small cell lung cancer (NSCLC). *Adv Clin Exp Med*. 2026;35(6):979–984. doi:10.17219/acem/210368

DOI
10.17219/acem/210368

Abstract

Background. Positron emission tomography (PET) using 18-fluorodeoxyglucose (^{18}-FDG) is the primary imaging modality for mediastinal staging of lung cancer. Higher standardized uptake value (SUV_{max}) reportedly correlates with more aggressive and more readily detectable tumors. However, high-quality scientific evidence regarding the correlation between primary tumor SUV_{max} and PET performance in detecting mediastinal lymph node metastases remains limited.

Objectives. To analyze the correlation between ^{18}-FDG SUV_{max} and the diagnostic yield of PET, as well as the impact of clinical factors such as age, sex, body mass index (BMI), primary tumor lobar location, and histological type on this relationship.

Materials and methods. This retrospective analysis utilized an institutional database and included a consecutive cohort of patients who underwent lung cancer surgery.

Results. In the overall cohort of 774 patients, PET demonstrated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of 59%, 76%, 27%, and 92%, respectively. Significant differences in sensitivity were observed between the $SUV_{max} < 10$ and $SUV_{max} > 15$ groups ($p = 0.007$), as well as between the SUV_{max} 10–15 and $SUV_{max} > 15$ groups ($p = 0.001$). Significant differences in NPV were found between the $SUV_{max} < 10$ and SUV_{max} 10–15 groups ($p = 0.011$). Logistic regression analysis revealed no association between the risk of false negative (FN) PET results for detecting N2 disease and patient characteristics, primary tumor lobar location, or histological type.

Conclusions. Our study confirms that higher SUV_{max} values of the primary tumor correlate with improved PET diagnostic performance in assessing mediastinal lymph node metastasis. A novel finding of this study is that clinical variables such as age, BMI, sex, tumor location, and histological type did not significantly influence the risk of FN results. The main limitations of this study are its retrospective design and single-institution cohort, which may not fully represent the diversity of patient demographics, disease characteristics, and treatment practices across different healthcare settings and populations.

Key words: lung neoplasms, neoplasm staging, lymph nodes, positron emission tomography, standardized uptake value

Highlights

- Limited data are available regarding the correlation between primary tumor SUV_{max} and positron emission tomography (PET) accuracy in detecting mediastinal lymph node metastases in non-small cell lung cancer (NSCLC).
- Higher SUV_{max} of the primary tumor is strongly associated with improved PET sensitivity for detecting mediastinal lymph node metastases in patients with NSCLC.
- PET sensitivity significantly increases in NSCLC tumors with $SUV_{max} > 15$ compared with tumors in lower SUV_{max} groups, thereby enhancing diagnostic performance.
- Clinical factors such as age, body mass index (BMI), sex, tumor location, and histology show no significant effect on false negative PET results in N2 lymph node assessment.

Background

Lung cancer is one of the leading causes of cancer-related mortality worldwide.[1,2] Positron emission tomography (PET) using 18-fluorodeoxyglucose (^{18}-FDG) is the primary imaging modality for mediastinal staging of lung cancer. In particular, the negative predictive value (NPV) of PET is high.[3–7] False positive PET results are common due to the glucose-dependent metabolism of leukocytes and granulation tissue. These mainly affect the specificity of PET and necessitate pathological confirmation of PET-positive lesions. However, as the primary goal of pretreatment evaluation is to exclude N2 disease, NPV remains the most important parameter.

Higher standardized uptake value (SUV_{max}) values reportedly correlate with more aggressive and more easily detectable tumors.[8,9] However, this does not directly translate into the risk of mediastinal lymph node involvement. Reliable assessment of the correlation between specific clinical characteristics and the probability of N2 disease requires analysis of large patient cohorts treated according to a uniform protocol. Such reports are lacking. Therefore, high-quality scientific data regarding the correlation between primary tumor SUV_{max} and PET performance in detecting mediastinal lymph node metastases remain limited. In addition, our understanding of the influence of different clinical variables on the diagnostic yield of PET is still incomplete.

Current guidelines from the European Society of Thoracic Surgeons (ESTS) indicate a low risk of N2 disease in cases of PET-negative mediastinum, tumor diameter <3 cm, and tumor location within the outer 1/3 of the lung. These guidelines do not consider primary tumor SUV_{max} as a risk factor for N2 involvement.[10] From a practical perspective, it is important to determine whether similar diagnostic performance of PET can be expected across sexes, body mass index (BMI) categories, primary tumor lobar locations, and histological cancer types.

Objectives

The primary objective of this study was to analyze the correlation between the SUV_{max} of ^{18}-FDG in primary tumors and the effectiveness of PET in diagnosing mediastinal lymph node involvement. The secondary objective was to evaluate the impact of clinical factors such as age, sex, BMI, primary tumor lobar location, and histological type on the rate of false negative (FN) PET results.

Materials and methods

Study design

This study employed a retrospective cohort design.

Clinical questions

1. Is there a correlation between the SUV_{max} of primary lung cancer and the diagnostic yield of PET in detecting mediastinal lymph node involvement?
2. What is the impact of specific clinical variables on the risk of obtaining FN results in N2 node assessment using PET?

Patients

The study included a consecutive group of surgically treated patients with non-small cell lung cancer (NSCLC). The inclusion criteria were: 1) age 18–90 years, 2) clinical stages I–IVA, and 3) curative-intent pulmonary resection with systematic lymph node dissection. The exclusion criteria were: 1) neoadjuvant treatment, 2) a final diagnosis other than NSCLC, and 3) incomplete mediastinal lymph node dissection.

Intervention

Positron emission tomography imaging was performed in all patients using a Discovery 690 scanner (GE Healthcare Chicago, USA). The upper limit of the allowed blood glucose concentration was 11 mmol/L. A weight-based formula for FDG dose calculation of 3–4 MBq/kg was used. The imaging protocol included computed tomography (CT) attenuation-correction imaging and lung

window reconstruction, with specific data acquisition parameters (80–210 mA, 3.75 mm section thickness, and 0.8-s gantry rotation speed). Whole-body PET scans were conducted with a 2.5 mm section thickness. Both non-attenuation-corrected (NAC) and measured attenuation-corrected (MAC) images were acquired using the Q.Clear algorithm (GE Healthcare). The interval between FDG injection and data acquisition was 45–60 min. The maximum SUV values were calculated using Pet Odyssey software (GE Healthcare), adjusted for patient weight and administered isotope dose. Positron emission tomography positivity was defined with reference to the SUV value of the aortic arch. The PET images were analyzed by the same team. A radiologist and a nuclear medicine specialist independently reviewed each PET scan with access to the clinical data.

Next, endobronchial ultrasound (EBUS) and endoscopic ultrasound (EUS) imaging of the mediastinum were performed with fine-needle biopsy of all PET-positive lymph nodes. Patients with no evidence of mediastinal spread were scheduled for curative-intent surgery. Pathological N stage, determined by examination of the surgical specimen and generally considered the gold standard, was used as the reference test. During surgery, anatomical lung resection and systematic lymph node dissection were performed according to the ESTS guidelines.[11] The consultant thoracic surgeon meticulously dissected lymph nodes from all stations and categorized them according to the 7th edition of the TNM classification.[8]

Postoperatively, tissue specimens were examined by an experienced lung pathologist with access to the clinical data. Standard light microscopy images with hematoxylin and eosin staining were used to diagnose nodal metastases (Olympus BX51; Olympus Corp., Tokyo, Japan).

Data extraction and handling

Relevant data, including demographics (age and sex), clinical details (cancer stage and histological type), treatment specifics (surgery type and PET scan parameters), and outcomes (pathological results and PET scan findings), were extracted from the hospital database. Extracted data were anonymized and securely stored to maintain patient confidentiality. Artificial intelligence (AI) was not used for the production of images or graphical elements in the paper, nor for data collection or analysis.

Statistical analyses

Statistica v. 13.5 PL (StatSoft Inc., Tulsa, USA) was used for statistical analysis. Logistic regression models were used to identify factors associated with FN PET results. Variables were grouped as follows: age (<60, 60–70, and >70 years), BMI (18.5–25 kg/m^2, >25–30 kg/m^2, and >30 kg/m^2), and SUV_{max} (<10, 10–15, and >15). Both univariate and multivariate models were constructed. Initially, the forward stepwise selection method was applied to identify promising predictor variables efficiently. The results were then compared with those obtained using best subset selection, which examines all possible predictor combinations but is more computationally intensive. Odds ratios (ORs) with 95% confidence intervals (95% CIs) calculated using the Clopper–Pearson exact method were used to interpret important factors. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated for the entire patient cohort and for subgroups based on sex, age, histological type, BMI, SUV_{max}, and primary tumor lobar location. The 95% CIs for sensitivity and specificity were also calculated and compared across specific SUV_{max} categories. Statistical significance was set at $p < 0.05$.

Results

In total, 884 patients were included. Of these, 110 met the exclusion criteria (89 received neoadjuvant chemotherapy, 17 had a final diagnosis other than NSCLC, and in 4 patients, complete mediastinal lymph node dissection was not performed due to intraoperative cardiopulmonary instability). Thus, data from 774 patients were analyzed. The characteristics of the study groups are presented in Table 1.

For the entire cohort, the sensitivity, specificity, PPV, and NPV of PET were 59%, 76%, 27%, and 92%, respectively. Detailed values according to sex, age, and BMI are presented in Table 2.

Separate analyses were performed for subgroups according to SUV_{max} categories: <10, 10–15, and >15. The results are summarized in Table 3.

A significant difference in sensitivity was found between the $SUV_{max} < 10$ and $SUV_{max} > 15$ groups ($p = 0.007$) and between the SUV_{max} 10–15 and $SUV_{max} > 15$ groups ($p = 0.000$). For NPV, a significant difference was found between the $SUV_{max} < 10$ and SUV_{max} 10–15 groups ($p = 0.011$). Logistic regression analysis did not show an association between the risk of FN PET results in detecting N2 disease and the following factors: age ($p = 0.769$), BMI ($p = 0.430$), sex ($p = 0.231$), primary tumor lobar location ($p = 0.312$–0.757), upper vs lower lobe location ($p = 0.126$), and histological type ($p = 0.623$–0.999). Simplified categorization into squamous and nonsquamous histology also showed no significant association ($p = 0.275$). Multivariate logistic regression analysis, which further investigated potential associations between FN results and clinical factors, also did not reveal significant relationships ($p = 0.079$–0.469).

Table 1. Characteristics of the study group

Variable		Value
Age [years]		30–87 (range) (65) (median)
Sex	F/M (n)	248/526
	F/M (%)	32/68
Tumor location, n (%)	RUL	209 (27)
	RML	31 (4)
	RLL	126 (16.3)
	RCE	56 (7.2)
	LUL	199 (25.7)
	LLL	115 (14.9)
	LCE	38 (4.9)
Histology, n (%)	ADC	234 (30.2)
	SCC	377 (48.7)
	ASC	63 (8.1)
	LCC	26 (3.5)
	Other	74 (9.5)
Stage, n (%)	IA	193 (25)
	IB	139 (17.9)
	IIA	49 (6.3)
	IIB	124 (16)
	IIIA	158 (20.5)
	IIIB	66 (8.5)
	IIIC	11 (1.4)
	IVA	34 (4.4)

ADC – adenocarcinoma; SCC – squamous cell carcinoma; ASC – adenosquamous carcinoma; LCC – large cell carcinoma; RCE – right central; LCE – left central; RUL – right upper lobe; RML – right middle lobe; RLL – right lower lobe; LUL – left upper lobe; LLL – left lower lobe.

Discussion

There is a large body of evidence demonstrating the strong prognostic value of N status.[12] The prognostic value of primary tumor SUV_{max} has been confirmed in several studies.[13–15] Our analysis of data from 774 patients revealed that the sensitivity, specificity, PPV, and NPV of PET were 59%, 76%, 27%, and 92%, respectively. These values are consistent with previous studies that also demonstrated a high NPV of PET in evaluating mediastinal lymph node metastasis.[3–7]

Our findings revealed significant differences in sensitivity among the primary tumor SUV_{max} categories of <10, 10–15, and >15. The group with SUV_{max} > 15 exhibited significantly higher sensitivity, suggesting that higher glucose metabolism in the primary tumor was associated with improved detectability of lymph node metastases using PET. This may be attributable to a higher likelihood of nodal metastasis arising from less differentiated tumor cell clones. These findings are consistent with previous research indicating that higher SUV_{max} values correlate with more aggressive and more easily detectable tumors.[8,9] According to some authors, higher primary tumor SUV is associated with a higher rate of FN results regarding N2 node metastasis.[5–7,16] Our previous study demonstrated a correlation between primary tumor SUV_{max} and histological grade, further supporting the present findings.[17]

Logistic regression analysis did not show a significant association between the risk of FN PET results and factors such as age, BMI, sex, lobar tumor location, or histological type. This suggests that the diagnostic yield of PET for detecting N2 disease is consistent across different patient

Table 2. Diagnostic yield of positron emission tomography (PET) in detecting N2 disease in the whole cohort

Variable	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
All patients	58.82% (48.64–68.48%)	75.60% (72.17–78.80%)	26.79% (22.87–31.10%)	92.36% (90.52–93.87%)
Male	59.02% (45.68–71.45%)	72.90% (68.62–76.89%)	22.22% (18.10–26.97%)	93.13% (90.90–94.85%)
Female	58.54% (42.11–73.68%)	81.64% (75.69–86.67%)	38.71% (30.04–48.18%)	90.86% (87.29–93.50%)
Age <60 years	79.31% (60.28–92.01%)	76.03% (68.27–82.70%)	39.66% (31.79–48.09%)	94.87% (90.02–97.43%)
Age 60–70 years	46.94% (32.53–61.73%)	77.78% (72.93–82.13%)	23.71% (17.83–30.80%)	90.88% (88.38–92.88%)
Age >70 years	58.33% (36.64–77.89%)	71.50% (64.58–77.75%)	20.29% (14.51–27.63%)	93.24% (89.50–95.72%)
BMI 18.5–25 kg/m^2	60.87% (45.37–74.91%)	71.77% (65.73–77.29%)	28.57% (22.77–35.18%)	90.82% (87.24–93.46%)
BMI 25–30 kg/m^2	59.46% (42.10–75.25%)	79.69% (74.30–84.40%)	29.33% (22.48–37.27%)	93.27% (90.33–95.37%)
BMI > 30 kg/m^2	52.94% (27.81–77.02%)	72.97% (65.06–79.94%)	18.37% (11.79–27.47%)	93.10% (88.98–95.76%)

PPV – positive predictive value; NPV – negative predictive value; 95% CI – 95% confidence interval.

Table 3. Diagnostic yield of positron emission tomography (PET) for detecting N2 disease in subgroups according to SUV_{max}

SUV_{max}	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SUV_{max} < 10	57.14% (37.18–75.54%)	85.33% (80.42–89.40%)	29.63% (21.42–39.41%)	94.85% (92.29–96.59%)
SUV_{max} 10–15	51.28% (34.78–67.58%)	74.75% (68.18–80.59%)	28.17% (21.03–36.61%)	88.82% (85.08–91.72%)
SUV_{max} > 15	68.57% (50.71–83.15%)	64.45% (57.59–70.91%)	24.24% (19.34–29.93%)	92.52% (88.24–95.32%)

95% CI – 95% confidence interval.

populations. To the best of our knowledge, this is the only study to analyze this relationship. Our findings are important from a practical perspective, as they demonstrate that the diagnostic performance of PET is similar across sexes, age groups, BMI categories, tumor locations, and histological tumor types. Moreover, these findings suggest that greater caution should be exercised when assessing N2 disease in patients with primary tumor SUV_{max} < 15, as sensitivity for detecting N2 disease is considerably lower in this group. The lower sensitivity of PET in these patients may indicate the need for more extensive use of invasive mediastinal lymph node diagnostics.

The strength of our study lies in its large cohort size, the largest published to date, and the use of standardized imaging and pathological protocols across all analyzed patients. Moreover, the distribution of histological cancer types, with 79% adenocarcinoma (AC) and squamous cell carcinoma (SCC), is representative of that commonly observed in thoracic surgical practice.

Limitations of the study

The main limitation of this study is its retrospective design. In addition, the study cohort consisted of patients from a single institution, which may not fully represent the diversity of patient demographics, disease characteristics, and treatment practices observed across different healthcare settings and populations. Further prospective multicenter studies with larger patient cohorts are warranted to confirm our findings and identify potential factors affecting PET performance.

Conclusions

Our study confirms that higher SUV_{max} of the primary tumor correlates with improved PET diagnostic performance in assessing mediastinal lymph node metastasis. However, when the primary tumor SUV_{max} values are below 15, N2 assessment should be approached with greater caution, as sensitivity for detecting N2 disease in this group is significantly lower. Positron emission tomography demonstrated a high NPV, making it valuable for ruling out N2 metastasis. Additionally, clinical variables such as age, BMI, sex, tumor location, and histological type did not significantly influence the risk of FN results. Further research is needed to better understand these phenomena and optimize imaging protocols, thereby improving clinical outcomes in patients with lung cancer.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.16977760. The package contains the following files:

Supplementary Table 1. Box–Tidwell test.
Supplementary Table 2. Variance inflation factor (VIF).
Supplementary Table 3. Cooks's distance.

Data Availability Statement

The datasets supporting the findings of this study are not publicly available due to ethical and privacy restrictions related to sensitive patient information. Although the data have been anonymized, they contain health information protected under privacy laws and unrestricted public sharing could risk patient confidentiality and lead to misuse by unauthorized third parties. The data are available from corresponding author upon reasonable request.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Adam Kużdżał https://orcid.org/0009-0008-9370-7967
Karolina Gambuś https://orcid.org/0000-0001-9675-1623
Konrad Moszczyński https://orcid.org/0009-0004-6012-6470
Sofiia Popovchenko https://orcid.org/0009-0000-0713-8713
Katarzyna Żanowska https://orcid.org/0000-0003-4863-681X
Piotr Kocoń https://orcid.org/0000-0002-4685-279X

References

1. Chaitanya Thandra K, Barsouk A, Saginala K, Sukumar Aluru J, Barsouk A. Epidemiology of lung cancer. *Wspolczesna Onkol*. 2021; 25(1):45–52. doi:10.5114/wo.2021.103829
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7–30. doi:10.3322/caac.21590

3. Farrell MA, McAdams HP, Herndon JE, Patz EF. Non-small cell lung cancer: FDG PET for nodal staging in patients with stage I disease. *Radiology*. 2000;215(3):886–890. doi:10.1148/radiology.215.3.r00jn29886
4. Ventura E, Islam T, Gee MS, Mahmood U, Braschi M, Harisinghani MG. Detection of nodal metastatic disease in patients with non-small cell lung cancer: Comparison of positron emission tomography (PET), contrast-enhanced computed tomography (CT), and combined PET-CT. *Clin Imaging*. 2010;34(1):20–28. doi:10.1016/j.clinimag.2009.03.012
5. Bustos García De Castro A, Ferreirós Domínguez J, Delgado Bolton R, et al. PET-CT in presurgical lymph node staging in non-small cell lung cancer: The importance of false-negative and false-positive findings. *Radiologia*. 2017;59(2):147–158. doi:10.1016/j.rx.2016.12.001
6. Kaseda K, Watanabe K, Asakura K, Kazama A, Ozawa Y. Identification of false-negative and false-positive diagnoses of lymph node metastases in non-small cell lung cancer patients staged by integrated ^{18F}-fluorodeoxyglucose-positron emission tomography/computed tomography: A retrospective cohort study. *Thorac Cancer*. 2016;7(4):473–480. doi:10.1111/1759-7714.12358
7. Li S, Zheng Q, Ma Y, et al. Implications of false negative and false positive diagnosis in lymph node staging of NSCLC by means of 18F-FDG PET/CT. *PLoS One*. 2013;8(10):e78552. doi:10.1371/journal.pone.0078552
8. Cengiz A, Aydın F, Sipahi M, et al. The role of F-18 FDG PET/CT in differentiating benign from malignant pulmonary masses and accompanying lymph nodes. *Tuberk Toraks*. 2018;66(2):130–135. PMID:30246656.
9. Takenaka T, Yano T, Morodomi Y, et al. Prediction of true-negative lymph node metastasis in clinical IA non-small cell lung cancer by measuring standardized uptake values on positron emission tomography. *Surg Today*. 2012;42(10):934–939. doi:10.1007/s00595-012-0277-7
10. De Leyn P, Dooms C, Kuzdzal J, et al. Revised ESTS guidelines for preoperative mediastinal lymph node staging for non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2014;45(5):787–798. doi:10.1093/ejcts/ezu028
11. Lardinois D, Deleyn P, Van Schil P, et al. ESTS guidelines for intraoperative lymph node staging in non-small cell lung cancer. *Eur J Cardiothorac Surg*. 2006;30(5):787–792. doi:10.1016/j.ejcts.2006.08.008
12. Asamura H, Chansky K, Crowley J, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project. *J Thorac Oncol*. 2015;10(12):1675–1684. doi:10.1097/JTO.0000000000000678
13. Imamura Y, Azuma K, Kurata S, et al. Prognostic value of SUV_{max} measurements obtained by FDG-PET in patients with non-small cell lung cancer receiving chemotherapy. *Lung Cancer*. 2011;71(1):49–54. doi:10.1016/j.lungcan.2010.04.004
14. Li Y, Wu X, Huang Y, Bian D, Jiang L. ^{18}F-FDG PET/CT in lung adenosquamous carcinoma and its correlation with clinicopathological features and prognosis. *Ann Nucl Med*. 2020;34(5):314–321. doi:10.1007/s12149-020-01450-1
15. Qiu X, Liang H, Zhong W, et al. Prognostic impact of maximum standardized uptake value on ^{18}F-FDG PET CT imaging of the primary lung lesion on survival in advanced non-small cell lung cancer: A retrospective study. *Thorac Cancer*. 2021;12(6):845–853. doi:10.1111/1759-7714.13863
16. Damirov F, Büsing K, Yavuz G, et al. Preoperative hilar and mediastinal lymph node staging in patients with suspected or diagnosed lung cancer: Accuracy of ^{18}F-FDG-PET/CT. A retrospective cohort study of 138 patients. *Diagnostics (Basel)*. 2023;13(3):403. doi:10.3390/diagnostics13030403
17. Kużdżał B, Moszczyński K, Żanowska K, et al. Correlation between 18-FDG standardized uptake value and tumor grade in patients with resectable non-small cell lung cancer. *Transl Cancer Res*. 2023;12(12):3530–3537. doi:10.21037/tcr-23-798

The role of static and dynamic evaluation of sarcopenia in liver transplant candidates

Maciej Miarka[1,A,B,D,E], Wiktor Smyk[2,D], Aleksandra Bodys-Pełka[3,B,C], Krzysztof Gibiński[4,C], Renata Główczyńska[3,A,C,D], Wojciech Figiel[5,B,C], Joanna Raszeja-Wyszomirska[1,A,D–F]

[1] Department of Hepatology, Transplantology and Internal Medicine, Medical University of Warsaw, Poland
[2] Department of Gastroenterology and Hepatology, Medical University of Gdańsk, Poland
[3] 1st Department of Cardiology, Medical University of Warsaw, Poland
[4] 2nd Department of Clinical Radiology, Medical University of Warsaw, Poland
[5] Department of General, Transplant and Liver Surgery, Medical University of Warsaw, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):985–995

Address for correspondence
Joanna Raszeja-Wyszomirska
E-mail: jorasz@gmail.com

Funding sources
None declared

Conflict of interest
None declared

Received on September 15, 2024
Reviewed on July 14, 2025
Accepted on October 13, 2025

Published online on May 5, 2026

Abstract

Background. Sarcopenia, characterized by the loss of skeletal muscle mass, strength, and function, is a prevalent and severe complication of liver cirrhosis, irrespective of its etiology. Despite its early onset in cirrhosis, the impact of sarcopenia on liver transplantation (LT) outcomes is often underestimated.

Objectives. This study evaluates the role of both static and dynamic assessments of sarcopenia in LT recipients, focusing on cardiopulmonary performance and the likelihood of prolonged ICU stays.

Materials and methods. We studied 54 LT recipients (median age: 53.5 years, 59% female patients) at a single center. The L3 skeletal muscle index (L3SMI) was measured using computed tomography (CT) scans, while exercise tolerance was evaluated with the 6-minute walk test (6MWT). Cardiac output (CO) was recorded in liters per minute.

Results. The median Model for End-Stage Liver Disease (MELD) score was 14.4, with 37% of patients classified as Child–Pugh class C. Major LT indications included autoimmune liver diseases (46.3%) and alcohol-related liver disease (ALD; 31.5%). Sarcopenia was present in 60.4% of patients. No significant differences in L3SMI were found related to underlying liver disease, gender, body mass index (BMI), or ammonia levels. However, patients with ALD covered significantly shorter distances in the 6MWT ($p = 0.02$). Sarcopenic patients had significantly higher CO than non-sarcopenic patients ($p < 0.01$). Intensive care unit (ICU) stay ≥3 days was observed in 77.8% of recipients, with no clear risk factors identified.

Conclusions. Sarcopenia linked to end-stage liver disease (ESLD) correlates with cirrhotic cardiomyopathy, as evidenced by increased CO. Further studies are required to clarify the role of 6MWT and CO in post-LT risk stratification.

Key words: liver transplantation, cardiopulmonary exercise test, sarcopenia, cardiac output, 6-minute walk test

Cite as
Miarka M, Smyk W, Bodys-Pełka A, et al. The role of static and dynamic evaluation of sarcopenia in liver transplant candidates. *Adv Clin Exp Med*. 2026;35(6):985–995. doi:10.17219/acem/212528

DOI
10.17219/acem/212528

Highlights

- Sarcopenia is highly prevalent among liver transplant (LT) recipients, with no significant correlation to liver function or ammonia levels.
- Patients with alcohol-related liver disease (ALD) showed the lowest 6-minute walk test (6MWT) results despite higher body mass index (BMI), indicating sarcopenic obesity.
- Increased cardiac output (CO) was observed in sarcopenic patients, suggesting a potential association with cirrhotic cardiomyopathy.
- The 6-minute walk test provided greater functional and clinical insight than the L3 skeletal muscle index (L3SMI) alone.
- Findings highlight the need for comprehensive pre-transplant evaluation, including cardiopulmonary fitness and assessment of comorbidities.

Background

Sarcopenia is a widely recognized, highly prevalent, and life-threatening complication in patients with liver cirrhosis, regardless of the disease's etiology. It is defined by a combination of generalized loss of skeletal muscle mass, strength, and function,[1] and it tends to develop earlier and more rapidly in liver cirrhosis compared with other chronic conditions. Alongside physical inactivity and frailty, sarcopenia is an independent predictor of morbidity and mortality in cirrhotic patients. Factors such as aging, malnutrition, decreased hepatic protein synthesis, catabolism, hyperammonemia, elevated pro-inflammatory cytokines, and low testosterone levels contribute to both cardiac and skeletal muscle deconditioning. This leads to diminished cardiac function and reserve, resulting in physical frailty and a reduction in skeletal muscle mass and strength.[2] As such, sarcopenia can be an indicator of poor general health when predicting outcomes in patients undergoing general surgery[3] and renal transplantation.[4] Male sex and liver function decompensation, as indicated by Child–Pugh class C cirrhosis, have been identified as independent predictors of sarcopenia, which in turn is an independent predictor of mortality.[5] Among liver transplant (LT) candidates, sarcopenia is linked to increased waitlist mortality[5–8] and sepsis-related deaths while on the waitlist.[8] Furthermore, sarcopenia tends to progress during the transplantation process and can persist up to 1 year post-surgery.[9] A recent meta-analysis highlighted the negative impact of sarcopenia on LT outcomes.[10] Post-transplant complications associated with sarcopenia include extended hospital and intensive care unit (ICU) stays, a higher risk of infections,[11] and increased mortality following transplantation.[12–17]

Computed tomography has been established as the gold standard for identifying low muscle mass, as recognized by the European Working Group on Sarcopenia in Older People, using measurements at the L3 vertebral level. However, the same Working Group also recommends dynamic evaluations of muscle function, which may provide a more objective assessment of tissue aging and natural remodeling. Recent data suggest that sarcopenia may impact the left heart chamber, contributing to occult heart failure with preserved ejection fraction under baseline conditions,[18] as well as leading to respiratory muscle and diaphragmatic sarcopenia and associated weakness.[19,20]

While outcomes following LT are influenced by recipient, donor, and procedural factors, the scarcity of donor livers necessitates optimizing the life-saving potential of procured livers, particularly in older candidates with more advanced end-stage liver disease (ESLD) and additional comorbidities (e.g., metabolic syndrome). Moreover, cardiovascular events are a significant contributor to serious complications and non-liver-related deaths after LT.[21] Pre-transplant cardiovascular assessments aim to detect severe cardiopulmonary diseases that could render LT futile. Typically, cardiological evaluations focus on ischemic heart disease and its complications. However, there is a noticeable lack of routine dynamic assessments of cardiopulmonary performance in ESLD candidates under conditions of increased oxygen demand or hemodynamic instability.

Exercise intolerance is widely regarded as an indicator of reduced functional capacity and disability. In ESLD, there is a complex interplay of factors such as malnutrition, sarcopenia, and frailty.[22] The 6-minute walk test (6MWT) is commonly employed to assess exercise capacity. It is an inexpensive, well-tolerated, noninvasive test that correlates with daily activity levels. Numerous studies have shown that 6MWT results correlate with the Model for End-Stage Liver Disease (MELD) and Child–Pugh scores as well as with mortality rates.[23,24] The 6MWT assesses overall exercise capacity and may correlate with maximal oxygen consumption (VO_2), predicting survival in cirrhotic patients and the risk of death while on the LT waiting list.[24] Poor performance during the 6MWT indicates abnormal exercise capacity in cirrhotic patients, with shorter distances covered correlating with more severe liver disease as assessed with the Child–Pugh score[24] and with short-term survival.[25] A peak oxygen uptake (VO_2 peak) below 60% of the predicted value has been identified as a prognostic factor for 100-day LT-related mortality.[26]

Conversely, liver cirrhosis is associated with increased resting cardiac output (CO), decreased systemic vascular resistance, systolic incompetence under stress, diastolic dysfunction, increased left ventricular thickness, and electrocardiography (ECG) abnormalities, collectively referred to as hyperdynamic circulation and cirrhotic cardiomyopathy.[27,28] The increase in CO is linearly related to the progression of portal hypertension.[29] Notably, elevated CO is disproportionate to oxygen consumption, indicating a decreased arteriovenous oxygen difference.[30]

To date, the influence of sarcopenia on cardiopulmonary efficiency before LT and on subsequent morbidity and mortality after LT, particularly in terms of prioritizing patients by transplant survival benefit, has not been thoroughly investigated. However, it is rational to prioritize patients to avoid short-term futility, defined as post-transplant 3-month mortality, and to ensure long-term survival exceeding 50% at 5 years post-LT.

Objectives

This study evaluates the role of both static (Skeletal Muscle Index at the 3rd lumbar vertebra (L3SMI)) and dynamic (6MWT) assessments of sarcopenia in LT recipients, focusing on cardiopulmonary performance, body composition differences across liver disease etiologies, and their impact on the likelihood of prolonged ICU stays. Additionally, this study aims to assess the relationship between sarcopenia, functional performance, and body composition, highlighting the significance of 6MWT as a functional marker and the influence of liver disease etiology on body mass index (BMI) and muscle mass.

Materials and methods

Participants

The study group consisted of 54 well-characterized liver graft recipients (22 men and 32 women), with a median age of 53.5 years. The liver disease etiology included alcohol-related liver disease (ALD; 31.5% of patients), autoimmune liver diseases (46.3%), viral hepatitis (13%), and other causes (9%). The median MELD score for the entire cohort was 14.4 points, with 37% of patients classified as Child–Pugh class C. Median BMI was 23.95 kg/m^2 for women and 27.65 kg/m^2 for male patients. L3SMI was measured using computed tomography (CT) scans, with a median value of 46.1 cm^2/m^2 for women and 44.55 cm^2/m^2 for men. The 6MWT was performed indoors along a flat, straight, 30-m corridor, with the median distance covered being 431 m. Cardiac output was measured using a noninvasive CNAP device during the 6MWT. Intensive care unit stay duration was recorded, with 77.8% of recipients staying more than 3 days in ICU and 9.5% staying more than 6 days. Two patients died within 1 year after transplantation. The characteristics of the study group are summarized in Table 1.

The study did not require an informed consent form to be signed. All the described tests are routinely performed in LT candidates. The scheme of data gathering and evaluation is presented in Fig. 1.

Definition of sarcopenia

Computed tomography scanning at the L3 vertebral level has been validated as the gold standard for recognizing low muscle mass, according to the European Working Group on Sarcopenia in Older People.[1] Following the method proposed by Carey et al.,[31] the cross-sectional area of muscles at L3 was measured and normalized to the individual's height, resulting in the skeletal muscle index (L3SMI). Cut-off values of 50 cm^2/m^2 for men and 39 cm^2/m^2 for women were used to define sarcopenia.[31] Although the L3SMI score was used to define sarcopenia, we acknowledge that current consensus definitions recommend incorporating both low muscle mass and reduced muscle strength or physical performance. In our cohort, direct measures of muscle strength, such as handgrip dynamometry, were not available due to logistical constraints. However, we employed the 6MWT as a surrogate marker of functional capacity (see below).

Table 1. Descriptive characteristics of liver transplant candidates grouped by liver disease etiology

Variable	ALD	Autoimmune	Other	Viral
Age [years], median (IQR)	59.0 (50.5–64.5)	49.0 (41.0–61.0)	60.0 (47.0–63.3)	58.5 (57.8–59.3)
MELD score [points], median (IQR)	15.0 (13.0–20.0)	12.0 (9.0–16.0)	12.0 (8.8–15.3)	19.5 (17.3–21.8)
BMI [kg/m^2], median (IQR)	27.1 (23.0–30.2)	21.8 (20.7–24.7)	30.8 (28.5–32.6)	22.6 (21.3–23.9)
L3SMI [cm^2/m^2], median (IQR)	43.1 (40.2–49.0)	46.1 (35.4–49.9)	49.6 (38.9–52.8)	47.2 (39.1–55.3)
6MWT [m], median (IQR)	375.0 (325.0–423.5)	450.0 (370.0–550.0)	467.5 (430.0–493.8)	432.0 (408.0–456.0)
CO [L/min], median (IQR)	5.8 (4.6–7.7)	5.4 (4.5–6.2)	6.2 (5.3–6.8)	6.0 (5.6–6.4)
ICU stay >3 days, n	13	15	6	1
ICU stay >6 days, n	3	1	0	0

Values are presented as median and interquartile range (IQR) unless otherwise indicated. MELD – Model for End-Stage Liver Disease; BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; ICU – intensive care unit; CO – cardiac output; ALD – alcohol-related liver disease.

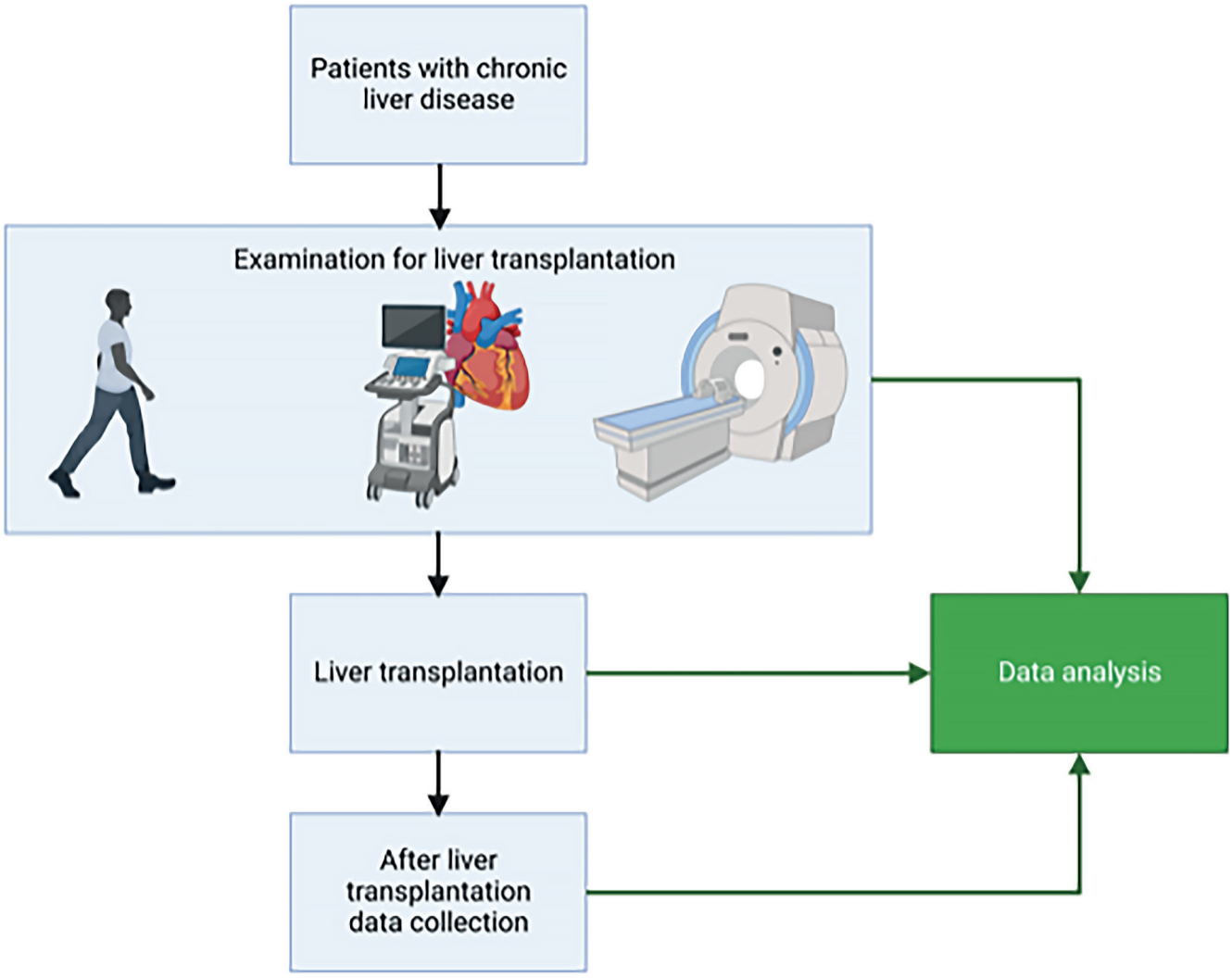


Fig. 1. The scheme of data gathering and evaluation

6-minute walk test and hemodynamic measurements

The 6MWT was conducted according to guidelines provided by the Polish Respiratory Society,[32] under the supervision of a qualified physician. The test was performed indoors, on a flat, straight, 30-m-long corridor with a hard surface and free of any type of obstacles. Patients were instructed to walk as fast as possible for 6 min, with the distance covered recorded at the end. Cardiac output was continuously measured during the test using a continuous noninvasive arterial pressure (CNAP) device for noninvasive, continuous measurement of hemodynamic parameters (CNSystems Medizintechnik AG, Graz, Austria) during the 6MWT.

Statistical analyses

Statistical analyses were performed using IBM SPSS v. 26.0 (IBM Corp., Armonk, USA), GraphPad Prism v. 8.0 (GraphPad Software, San Diego, USA), and Python (StatsModels package) for logistic regression modeling. A two-sided p-value of less than 0.05 was considered statistically significant.

The distribution of continuous variables was assessed using the Shapiro–Wilk test. Variables with normal distribution were compared between 2 groups using the Student's t-test. For non-normally distributed variables, the Mann–Whitney U test was applied. Comparisons involving more than 2 groups were conducted using the Kruskal–Wallis test. Homogeneity of variances was evaluated with Levene's test, and results are presented in Supplementary Table 1.

Correlations between continuous variables were assessed using Spearman's rank correlation coefficient. Univariate logistic regression was used to evaluate associations between clinical and functional parameters and prolonged ICU stay (>3 days and >6 days). Results are reported as odds ratios (OR) with 95% confidence intervals (95% CIs), p-values, and model performance indices. Model fit was assessed using the Nagelkerke R^2 statistic, which is reported in Tables 2,3.

To verify the assumption of linearity between continuous predictors and the logit of the binary outcome

Table 2. Univariate logistic regression analysis of risk factors for prolonged ICU stay (>3 days) in liver transplant recipients

Factor	OR (95% CI)	p-value	Nagelkerke R^2
6MWT (per 1 m increase)	0.991 (0.981–0.999)	0.047	0.034
L3SMI (per 1 cm^2/m^2 increase)	0.970 (0.927–0.998)	0.049	0.036
BMI (per 1 kg/m^2 increase)	0.927 (0.842–1.021)	0.104	0.053
CO (per 1 L/min increase)	0.359 (0.163–0.766)	0.010	0.008
MELD score (per 1 point increase)	1.089 (0.982–1.211)	0.105	0.013

Results are presented as odds ratios (ORs) with 95% confidence intervals (95% CIs) and corresponding p-values. MELD – Model for End-Stage Liver Disease; BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; ICU – intensive care unit; CO – cardiac output.

Table 3. Univariate logistic regression analysis of risk factors for extended ICU stay (>6 days) in liver transplant recipients

Factor	OR (95% CI)	p-value	Nagelkerke R^2
6MWT (per 1 m increase)	0.986 (0.969–0.997)	0.025	0.019
L3SMI (per 1 cm^2/m^2 increase)	0.962 (0.910–0.999)	0.048	0.157
BMI (per 1 kg/m^2 increase)	0.873 (0.748–0.991)	0.046	0.005
CO (per 1 L/min increase)	0.174 (0.057–0.529)	0.003	0.013
MELD score (per 1 point increase)	1.148 (1.002–1.363)	0.048	0.032

Results are presented as odds ratios (ORs) with 95% confidence intervals (95% CIs) and corresponding p-values. MELD – Model for End-Stage Liver Disease; BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; ICU – intensive care unit; CO – cardiac output.

Table 4. Association between sarcopenia-related parameters and demographic/clinical variables in liver transplant recipients

Variable	L3SMI [cm^2/m^2] (p-value)	6MWT [m] (p-value)	CO [L/min] (p-value)	Statistical test
Age (<50 vs >50 years)	0.950	0.410	0.003	Mann–Whitney U
Sex (man vs women)	0.620	0.160	<0.001	Mann–Whitney U
Etiology (ALD vs viral vs autoimmune)	0.840	0.060	0.290	Kruskal–Wallis
BMI [kg/m^2]	0.160	0.090	0.020	Mann–Whitney U
Ammonia level [mg/dL]	0.690	0.160	0.580	Mann–Whitney U

The p-values are derived from Mann–Whitney U test for binary comparisons and Kruskal–Wallis test for comparisons across more than 2 groups. BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; CO – cardiac output; ALD – alcohol-related liver disease.

in logistic regression, the Box–Tidwell test was employed. This method introduces interaction terms between each predictor and its natural logarithm. Nonsignificant interaction terms ($p > 0.33$ for all tested variables) confirmed that the linearity assumption was met for all continuous predictors included in the model.

Results

A total of 54 patients undergoing LT were included in the study. Of these, 32 (60.4%) met the criteria for sarcopenia based on L3SMI thresholds. Sarcopenia was more common among male recipients, with 26 out of 32 male patients (81.3%) classified as sarcopenic (L3SMI < 50 cm^2/m^2), compared to 6 out of 22 women (27.3%) with L3SMI < 39 cm^2/m^2. The lowest recorded L3SMI in the male group was 24.8 cm^2/m^2 (Table 1).

No statistically significant associations were found between sarcopenia and age (≤50 vs >50 years, $p = 0.95$), gender ($p = 0.62$), BMI ($p = 0.16$), liver disease etiology ($p = 0.84$), ammonia levels ($p = 0.69$), or liver function as measured using the MELD score ($p = 0.73$). Similar patterns were observed for 6MWT performance and CO. Detailed comparisons are presented in Table 4.

Etiology-specific analysis revealed significant differences in BMI ($p = 0.004$, Kruskal–Wallis test). Post hoc Bonferroni-adjusted Mann–Whitney U tests showed that ALD patients had significantly higher BMI compared to autoimmune ($p = 0.007$) and viral ($p = 0.004$) subgroups. Median BMI was 27.1 kg/m^2 (ALD), 21.75 kg/m^2 (autoimmune), and 22.58 kg/m^2 (viral). The MELD scores were numerically highest among viral hepatitis patients (median: 19.5), but this did not reach statistical significance ($p = 0.07$).

No significant differences in L3SMI ($p = 0.84$) or CO ($p = 0.29$) were observed across liver disease etiologies. The 6MWT performance was lowest in ALD patients (median: 375.0 m), and highest in other etiologies (up to 467.5 m), with borderline significance ($p = 0.06$). As shown in Fig. 2, histograms with kernel density estimates were used to assess the distribution of continuous variables. This evaluation informed the selection of appropriate statistical methods (parametric vs nonparametric)

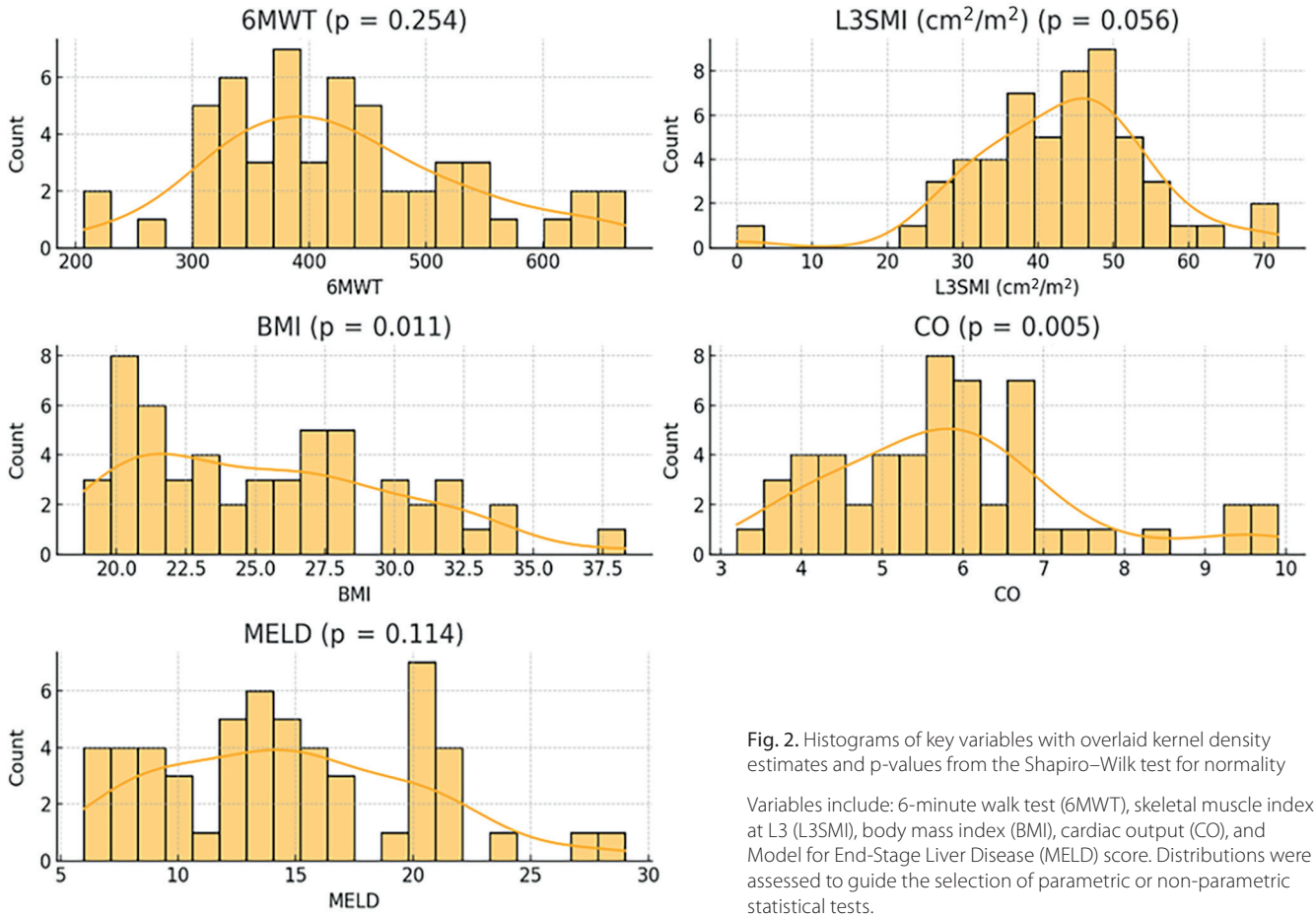


Fig. 2. Histograms of key variables with overlaid kernel density estimates and p-values from the Shapiro–Wilk test for normality

Variables include: 6-minute walk test (6MWT), skeletal muscle index at L3 (L3SMI), body mass index (BMI), cardiac output (CO), and Model for End-Stage Liver Disease (MELD) score. Distributions were assessed to guide the selection of parametric or non-parametric statistical tests.

for further analyses. Levene's test results for homogeneity of variance are provided in Supplementary Table 1.

Table 5 summarizes descriptive statistics for clinical and functional measures. Spearman correlations between parameters are presented in Table 6. L3SMI was significantly associated with BMI ($p < 0.05$), and 6MWT correlated positively with both L3SMI and CO ($p < 0.01$).

Univariate logistic regression models evaluating risk factors for ICU stay >3 days are summarized in Table 2. None of the predictors reached statistical significance. However, for ICU stay >6 days (Table 6), lower values of 6MWT ($p = 0.025$), L3SMI ($p = 0.048$), BMI ($p = 0.046$), and CO ($p = 0.003$), as well as higher MELD scores ($p = 0.048$), were associated with prolonged ICU stay. Nagelkerke R^2 values ranged from 0.005 to 0.157.

Table 5. Summary of clinical and functional characteristics of the study population

Variable	Median (IQR)
6MWT [m]	380.0 (300.0–450.0)
L3SMI [cm²/m²]	44.1 (39.5–49.2)
BMI [kg/m²]	25.7 (22.3–28.4)
CO [L/min]	5.8 (4.7–6.5)
MELD score [points]	15.0 (12.0–18.0)

Values are presented as median with interquartile range (IQR). MELD – Model for End-Stage Liver Disease; BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; ICU – intensive care unit; CO – cardiac output.

To formally assess the assumption of linearity between continuous predictors and the logit of the binary outcome (prolonged ICU stay >6 days), the Box–Tidwell test was conducted for the following variables: age, L3SMI, total cross-sectional muscle area at L3, and 6MWT distance. Each predictor was tested by including an interaction term between the variable and its natural logarithm in a logistic regression model.

None of the interaction terms reached statistical significance (all $p > 0.33$), indicating that the assumption of linearity was met for all included continuous variables. Specifically, p-values for the interaction terms were as follows: age ($p = 0.359$), L3SMI ($p = 0.333$), total muscle area ($p = 0.351$), and 6MWT ($p = 0.351$). These results confirm that the logit-linear relationship required for logistic regression holds in this dataset.

Figure 3 presents boxplots of continuous clinical predictors stratified by ICU stay duration (≤6 days vs >6 days), with individual data points overlaid to illustrate intergroup variability. This visual representation complements the Box–Tidwell test results and provides an intuitive summary of the association between predictor variables and ICU outcome. Notably, patients with ICU stays longer than

Table 6. Spearman's correlation matrix between clinical and functional parameters in liver transplant recipients

Variable	MELD	BMI	CO	6MWT	L3SMI
MELD score [points]	1.00	–0.15	0.12	–0.22	–0.18
BMI [kg/m^2]	–0.15	1.00	–0.05	0.10	0.34*
CO [L/min]	0.12	–0.05	1.00	0.51**	0.33*
6MWT [m]	–0.22	0.10	0.51**	1.00	0.45**
L3SMI [cm^2/m^2]	–0.18	0.34*	0.33*	0.45**	1.00

*p < 0.05' ** p < 0.01. MELD – Model for End-Stage Liver Disease; BMI – body mass index; L3SMI – Skeletal Muscle Index at the 3rd lumbar vertebra; 6MWT – 6-minute walk test; ICU – intensive care unit; CO – cardiac output.

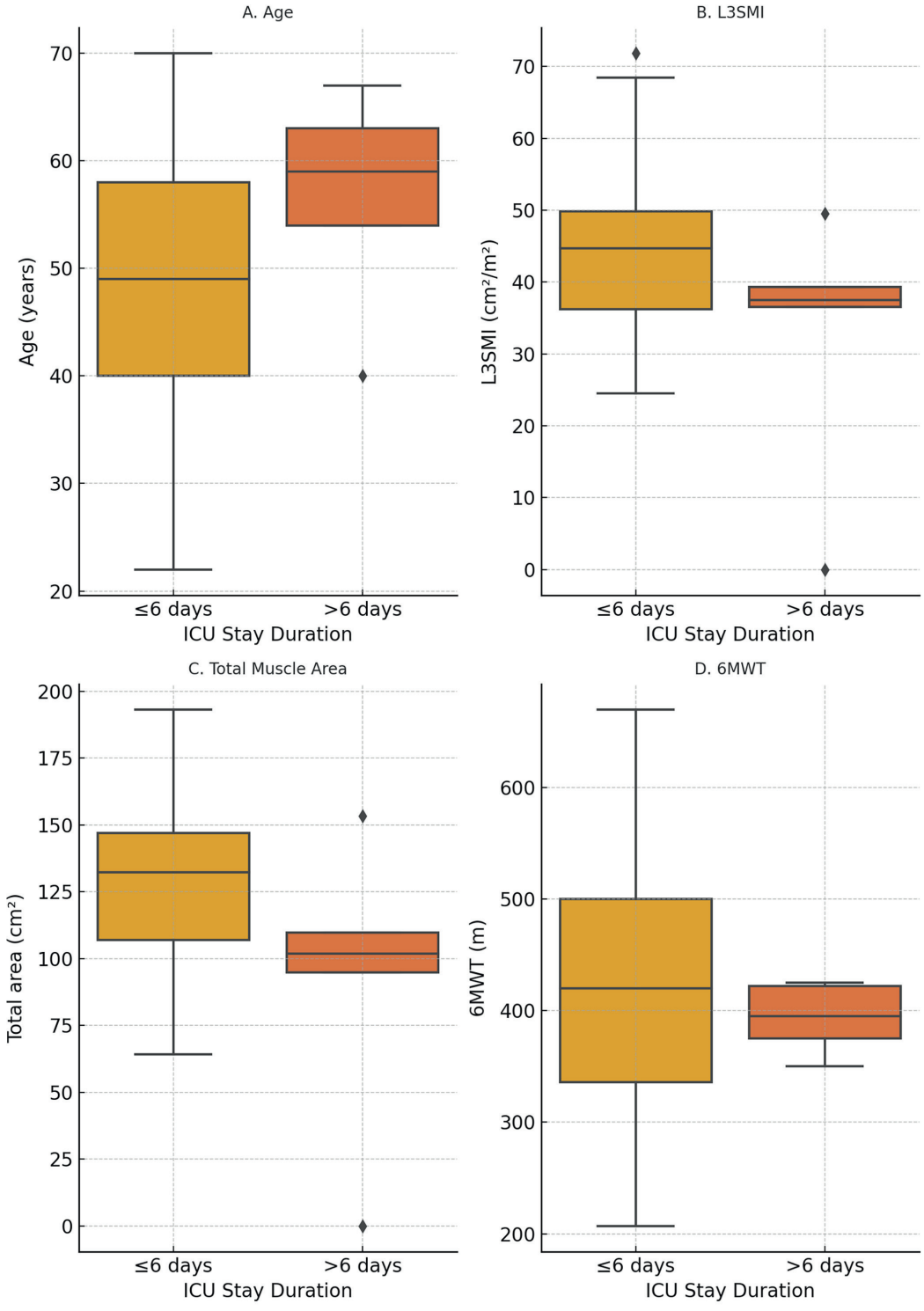


Fig. 3. Comparison of continuous predictors by intensive care unit (ICU) stay duration

Boxplots showing the distribution of (A) age, (B) L3 skeletal muscle index (L3SMI), (C) total cross-sectional muscle area at L3, and (D) 6-minute walk test (6MWT) distance, stratified by ICU stay duration (≤6 days vs >6 days). Individual data points are overlaid to show variability within groups. These plots provide visual context for the Box–Tidwell test results, which confirmed linearity between each predictor and the logit of the outcome variable

6 days exhibited lower values of L3SMI, 6MWT distance, and CO, supporting the hypothesis that reduced muscle mass and impaired functional capacity may contribute to prolonged postoperative recovery.

Discussion

Current cardiological assessment protocols for LT candidates primarily focus on detecting coronary artery disease and its associated complications. However, these protocols often fall short in evaluating patients who cannot undergo cardiopulmonary exercise testing, particularly in the context of cirrhosis-related hemodynamic disturbances. Changes in lean muscle mass are closely linked to the pathophysiology of cardiovascular disease, perpetuating a cycle of functional decline.[33] This highlights a significant gap in assessing cardiac function during surgery, especially under conditions of increased oxygen demand and hemodynamic instability in LT recipients.

This study aimed to compare static (L3SMI) and dynamic (6MWT) sarcopenia assessments in LT recipients, while also evaluating CO as a potential indicator of cardiovascular fitness. Our prior research identified sarcopenia as a potential risk factor for prolonged ICU stays in LT candidates.[34] However, the current findings suggest that L3SMI alone does not adequately predict liver function, morbidity, or mortality post-transplantation. No significant correlation was observed between L3SMI-defined sarcopenia and liver disease etiology, BMI, ammonia levels, or functional capacity as measured with the 6MWT. This suggests that sarcopenia is highly prevalent among LT recipients, particularly men, but may not serve as a standalone risk marker for postoperative outcomes.

The visual comparison presented in Fig. 3 illustrates that patients with reduced 6MWT distances, lower L3SMI values, and diminished CO were more likely to experience prolonged ICU stays. Although no single variable independently predicted ICU duration, this consistent pattern suggests a multifactorial interplay between impaired muscular status, limited functional reserve, and reduced cardiovascular adaptability. Prolonged ICU stay is a clinically meaningful postoperative outcome, often associated with increased morbidity, resource utilization, and delayed recovery. These findings reinforce the importance of incorporating dynamic functional assessments into preoperative risk stratification, particularly in the context of post-transplant recovery, rather than relying solely on static anatomical metrics. Similar conclusions have been drawn in other transplant and surgical populations. For example, Singer et al. demonstrated that physical frailty phenotypes were strongly predictive of mortality after lung transplantation, highlighting the prognostic value of functional status.[35] In a broader perioperative context, Gillis et al. emphasized the role of prehabilitation in improving surgical outcomes and reducing postoperative complications.[36] In LT specifically, reduced 6MWT performance has been associated with mortality and adverse outcomes,[24] and physical frailty has been linked to prolonged ICU and hospital stays.[37] These data support integrating functional capacity testing, such as the 6MWT, into standard pre-transplant evaluation protocols.

A novel aspect of this study was the analysis of BMI variations across different liver disease etiologies. Alcohol-related liver disease patients exhibited significantly higher BMI values compared to autoimmune and viral liver disease groups ($p = 0.004$), yet they also demonstrated the poorest 6MWT performance. This aligns with prior research indicating that ALD is often associated with sarcopenic obesity, a condition characterized by higher fat deposition despite significant muscle wasting. In contrast, patients with autoimmune and viral liver diseases tended to have lower BMI values, potentially reflecting chronic inflammation, malabsorption, or metabolic deficiencies. These findings reinforce the need for individualized nutritional and metabolic assessments in LT candidates, as BMI alone does not fully capture muscle function or frailty risk.

From a pathophysiological perspective, advanced cirrhosis leads to compensatory cardiovascular adaptations, including a reduction in effective circulating blood volume, splanchnic vasodilation, and hyperdynamic circulation. This can manifest as increased resting CO and, in some cases, the development of cirrhotic cardiomyopathy (CCM). Typically, CCM remains subclinical, as compensatory mechanisms maintain myocardial function despite reduced preload and afterload. However, when exposed to stressors such as exercise or surgical interventions, patients with cirrhosis may exhibit impaired cardiovascular responses due to an inadequate increase in CO during exertion.[38,39] Despite this well-documented phenomenon, our findings did not reveal a significant association between sarcopenia status and CO. At the same time, CO was significantly associated with age, sex, and BMI, and positively correlated with L3SMI, indicating that its determinants may be multifactorial.

In healthy individuals, CO is expected to rise during exercise, driven by an increase in stroke volume and heart rate. However, in cirrhotic patients, dysautonomia, neurohormonal imbalances, and structural myocardial changes may blunt this response, leading to an inability to meet the metabolic demands of exercise.[40–42] This may partially explain why ALD patients, despite having a higher BMI, exhibited the poorest 6MWT results. Notably, patients with ALD covered significantly shorter distances than other etiological groups ($p = 0.02$), reinforcing the role of functional capacity assessments in this population.

While prolonged ICU stays in sarcopenic individuals have been linked to reduced respiratory efficiency and muscle dysfunction,[43] the present study did not identify specific predictors of postoperative ICU length of stay. Recent research suggests that sarcopenia may be associated with cardiac dysfunction, including left ventricular

impairment and heart failure with preserved ejection fraction, even under baseline conditions.[18] Additionally, respiratory muscle weakness and diaphragmatic dysfunction may exacerbate frailty and functional decline in cirrhotic patients.[19,20] A related condition, cardiac cachexia, has also been described in LT candidates, characterized by chronic inflammation, endotoxemia, and proteolytic activation, which accelerate muscle depletion and metabolic inefficiency.[43] This process is exacerbated by hyperammonemia, proteasomal overactivation, mitochondrial dysfunction, and increased oxidative stress, all of which contribute to skeletal muscle atrophy and exercise intolerance.[44,45]

Despite prior evidence suggesting an interplay between ammonia levels, sarcopenia, and cardiovascular dysfunction, this study did not find a direct correlation between ammonia and CO, 6MWT performance, or prolonged ICU stay. Additionally, L3SMI did not correlate with muscle function or transplant outcomes, reinforcing the limitations of static muscle assessments in cirrhotic patients. Notably, ALD patients had the lowest 6MWT results (median: 380 m), aligning closely with the mortality threshold of 387 m identified by Pimentel et al.[24] These findings underscore the clinical significance of exercise prehabilitation programs, which may improve functional capacity and post-transplant recovery.

Given that 6MWT performance provided more relevant functional insights than L3SMI, our findings suggest that functional rather than static muscle assessments may be more valuable in LT candidates. The lack of correlation between L3SMI and post-transplant outcomes raises important questions about whether alternative markers of sarcopenia, such as direct muscle strength measurements, should be prioritized in future studies. Moreover, BMI variations among liver disease etiologies suggest that body composition, rather than weight alone, may play a critical role in sarcopenia risk stratification. Future research should explore comprehensive body composition analyses, functional testing protocols, and prehabilitation interventions to optimize transplant candidate selection and postoperative outcomes.

Limitations of the study

A major limitation of this study is the relatively small sample size, which may reduce statistical power and limit the generalizability of the findings. However, the cohort reflects a real-world population undergoing a complex and infrequently performed procedure at a single tertiary center, which inherently limits recruitment capacity. Despite this, the prospective design, in-depth patient characterization, and combined functional and morphological assessment of sarcopenia provide meaningful preliminary insights and emphasize the need for future multicenter validation.

A key limitation of our study is the lack of detailed data on comorbidities such as cardiovascular disease, diabetes, and other metabolic disorders. In addition, renal function markers and hemoglobin levels were not included in the analysis, despite their potential influence on exercise tolerance and 6MWT performance. However, it is important to note that sarcopenia in the context of liver cirrhosis often develops earlier and progresses more rapidly than in other chronic conditions such as heart or lung diseases, particularly in subgroups such as those with ALD. In this population, sarcopenia is primarily driven by liver dysfunction itself, which may partially mitigate the impact of unmeasured comorbidities. Moreover, all patients in our cohort were functionally independent and able to complete the 6MWT, suggesting a relatively preserved baseline physical capacity across the group.

Although ammonia levels were evaluated, no significant associations were found with muscle strength, 6MWT results, or ICU length of stay. This may indicate that functional consequences of hyperammonemia-induced sarcopenia develop over longer timeframes, or that other mechanisms, such as chronic inflammation or neurohormonal dysregulation, play a more prominent role in sarcopenia related to liver cirrhosis.

Furthermore, the study did not identify clear predictors of prolonged ICU stay or early post-transplant complications, especially in relation to CO and pulmonary function. The elevated CO observed in sarcopenic patients raises important questions: Is this a compensatory response, a marker of advanced liver dysfunction, or an indication of cirrhotic cardiomyopathy? The current data are insufficient to resolve this issue and underscore the complex interplay between hepatic, muscular, and cardiovascular systems in this population.

Lastly, the lack of standardized pre-transplant functional evaluations remains a broader challenge in transplant candidate assessment. The significant differences in 6MWT performance between etiological subgroups, particularly the lower performance in ALD patients, highlight the potential value of incorporating objective cardiopulmonary fitness measures into routine transplant evaluations. Future research should prioritize multicenter, prospective studies with larger sample sizes and a broader clinical scope, including comorbidities, renal function, and inflammatory markers, to enhance risk stratification and improve outcomes in LT.

Conclusions

Our findings confirm that severe chronic liver insufficiency accompanied by sarcopenia is associated with cirrhotic cardiomyopathy, as evidenced by increased CO. However, rather than a simple linear relationship, these results suggest a complex, multifactorial interplay between sarcopenia, functional capacity, and hemodynamic adaptation in cirrhosis.

Importantly, ALD patients exhibited the lowest 6MWT performance despite having higher BMI, reinforcing the need for body composition analysis beyond traditional anthropometric measures. The lack of a significant correlation between L3SMI and transplant outcomes raises questions about whether static muscle mass assessment alone is sufficient for risk stratification in LT candidates. Instead, dynamic functional tests such as the 6MWT may provide more clinically relevant insights.

Further large-scale, multicenter studies are warranted to determine the precise impact of sarcopenia on cardiopulmonary performance in LT candidates, as well as to evaluate the role of early exercise prehabilitation programs in improving functional outcomes, reducing post-transplant complications, and optimizing long-term survival.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17494303. The package contains the following files:

Supplementary Table 1. Results of Levene's test for homogeneity of variances for key continuous variables.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo at https://doi.org/10.5281/zenodo.13851662.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Maciej Miarka https://orcid.org/0000-0002-5006-1077
Wiktor Smyk https://orcid.org/0000-0002-6767-2800
Aleksandra Bodys-Pełka https://orcid.org/0000-0003-1836-6425
Krzysztof Gibiński https://orcid.org/0000-0001-8763-8063
Renata Główczyńska https://orcid.org/0000-0003-1836-6425
Wojciech Figiel https://orcid.org/0000-0001-9716-7824
Joanna Raszeja-Wyszomirska https://orcid.org/0000-0001-7204-9784

References

1. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European consensus on definition and diagnosis. *Age Ageing*. 2010;39(4):412–423. doi:10.1093/ageing/afq034
2. Tandon P, Ismond KP, Riess K, et al. Exercise in cirrhosis: Translating evidence and experience to practice. *J Hepatol*. 2018;69(5):1164–1177. doi: 10.1016/j.jhep.2018.06.017
3. Englesbe MJ, Lee JS, He K, et al. Analytic morphomics, core muscle size, and surgical outcomes. *Ann Surg*. 2012;256(2):255–261. doi:10.1097/SLA.0b013e31826028b1
4. Streja E, Molnar MZ, Kovesdy CP, et al. Associations of pretransplant weight and muscle mass with mortality in renal transplant recipients. *Clin J Am Soc Nephrol*. 2011;6(6):1463–1473. doi:10.2215/CJN.09131010
5. Tandon P, Ney M, Irwin I, et al. Severe muscle depletion in patients on the liver transplant wait list: Its prevalence and independent prognostic value. *Liver Transpl*. 2012;18(10):1209–1216. doi:10.1002/lt.23495
6. Durand F, Buyse S, Francoz C, et al. Prognostic value of muscle atrophy in cirrhosis using psoas muscle thickness on computed tomography. *J Hepatol*. 2014;60(6):1151–1157. doi: 10.1016/j.jhep.2014.02.026
7. Meza-Junco J, Montano-Loza AJ, Baracos VE, et al. Sarcopenia as a prognostic index of nutritional status in concurrent cirrhosis and hepatocellular carcinoma. *J Clin Gastroenterol*. 2013;47(10):861–870. doi:10.1097/MCG.0b013e318293a825
8. Montano-Loza AJ, Meza-Junco J, Prado CMM, et al. Muscle wasting is associated with mortality in patients with cirrhosis. *Clin Gastroenterol Hepatol*. 2012;10(2):166–173.e1. doi: 10.1016/j.cgh.2011.08.028
9. Bhanji RA, Takahashi N, Moynagh MR, et al. The evolution and impact of sarcopenia pre- and post-liver transplantation. *Aliment Pharmacol Ther*. 2019;49(6):807–813. doi:10.1111/apt.15161
10. Van Vugt JLA, Levolger S, De Bruin RWF, Van Rosmalen J, Metselaar HJ, Izermans JNM. Systematic review and meta-analysis of the impact of computed tomography-assessed skeletal muscle mass on outcome in patients awaiting or undergoing liver transplantation. *Am J Transplant*. 2016;16(8):2277–2292. doi:10.1111/ajt.13732
11. Krell RW, Kaul DR, Martin AR, et al. Association between sarcopenia and the risk of serious infection among adults undergoing liver transplantation. *Liver Transpl*. 2013;19(12):1396–1402. doi:10.1002/lt.23752
12. DiMartini A, Cruz RJ, Dew MA, et al. Muscle mass predicts outcomes following liver transplantation. *Liver Transpl*. 2013;19(11):1172–1180. doi:10.1002/lt.23724
13. Englesbe MJ, Patel SP, He K, et al. Sarcopenia and mortality after liver transplantation. *J Am Coll Surg*. 2010;211(2):271–278. doi: 10.1016/j.jamcollsurg.2010.03.039
14. Hamaguchi Y, Kaido T, Okumura S, et al. Impact of quality as well as quantity of skeletal muscle on outcomes after liver transplantation. *Liver Transpl*. 2014;20(11):1413–1419. doi:10.1002/lt.23970
15. Kalafateli M, Mantzoukis K, Choi Yau Y, et al. Malnutrition and sarcopenia predict post-liver transplantation outcomes independently of the Model for End-stage Liver Disease score. *J Cachexia Sarcopenia Muscle*. 2017;8(1):113–121. doi:10.1002/jcsm.12095
16. Masuda T, Shirabe K, Ikegami T, et al. Sarcopenia is a prognostic factor in living donor liver transplantation. *Liver Transpl*. 2014;20(4):401–407. doi:10.1002/lt.23811
17. Montano-Loza AJ, Meza-Junco J, Baracos VE, et al. Severe muscle depletion predicts postoperative length of stay but is not associated with survival after liver transplantation. *Liver Transpl*. 2014;20(6):640–648. doi:10.1002/lt.23863
18. Kazemi-Bajestani SMR, Becher H, Ghosh S, Montano-Loza AJ, Baracos VE. Concurrent depletion of skeletal muscle, fat, and left ventricular mass in patients with cirrhosis of the liver: Letter to the Editor. *J Cachexia Sarcopenia Muscle*. 2016;7(1):97–99. doi:10.1002/jcsm.12093
19. Elliott JE, Greising SM, Mantilla CB, Sieck GC. Functional impact of sarcopenia in respiratory muscles. *Respir Physiol Neurobiol*. 2016; 226:137–146. doi: 10.1016/j.resp.2015.10.001
20. Yamada K, Kinugasa Y, Sota T, et al. Inspiratory muscle weakness is associated with exercise intolerance in patients with heart failure with preserved ejection fraction: A preliminary study. *J Cardiac Fail*. 2016;22(1):38–47. doi: 10.1016/j.cardfail.2015.10.010
21. D'Avola D, Cuervas-Mons V, Martí J, et al. Cardiovascular morbidity and mortality after liver transplantation: The protective role of mycophenolate mofetil. *Liver Transpl*. 2017;23(4):498–509. doi:10.1002/lt.24738
22. Saraswat VA, Kumar K. Untangling the web of malnutrition, sarcopenia, and frailty in chronic liver disease. *J Clin Exp Hepatol*. 2022;12(2):268–271. doi: 10.1016/j.jceh.2022.02.002
23. Agarwala P, Salzman SH. Six-minute walk test. *Chest*. 2020;157(3):603–611. doi: 10.1016/j.chest.2019.10.014
24. Pimentel CFMG, Amaral ACDC, Gonzalez AM, et al. Six-minute walking test performance is associated with survival in cirrhotic patients. *World J Hepatol*. 2021;13(11):1791–1801. doi:10.4254/wjh.v13.i11.1791
25. Cahalin LP, Mathier MA, Semigran MJ, Dec GW, DiSalvo TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. *Chest*. 1996;110(2):325–332. doi:10.1378/chest.110.2.325

26. Epstein SK, Freeman RB, Khayat A, Unterborn JN, Pratt DS, Kaplan MM. Aerobic capacity is associated with 100-day outcome after hepatic transplantation. *Liver Transpl*. 2004;10(3):418–424. doi:10.1002/lt.20088
27. Bodys-Pełka A, Kusztal M, Raszeja-Wyszomirska J, Główczyńska R, Grabowski M. What's new in cirrhotic cardiomyopathy? Review article. *J Pers Med*. 2021;11(12):1285. doi:10.3390/jpm11121285
28. Rimbaș RC, Rimbas M, Chitroceanu AM, Luchian LM, Pop C, Vinereanu D. Cirrhotic cardiomyopathy in the era of liver transplantation: Time for precise stepwise evaluation. *J Gastrointest Liver Dis*. 2020;29(4): 665–675. doi:10.15403/jgld-3137
29. Xanthopoulos A, Starling RC, Kitai T, Triposkiadis F. Heart failure and liver disease. *JACC Heart Fail*. 2019;7(2):87–97. doi: 10.1016/j.jchf.2018.10.007
30. Kowalski HJ, Abelmann WH. The cardiac output at rest in Laennec's cirrhosis. *J Clin Invest*. 1953;32(10):1025–1033. doi:10.1172/JCI102813
31. Carey EJ, Lai JC, Wang CW, et al. A multicenter study to define sarcopenia in patients with end-stage liver disease. *Liver Transpl*. 2017; 23(5):625–633. doi:10.1002/lt.24750
32. Przybyłowski T, Tomalak W, Siergiejko Z, et al. Polish Respiratory Society Guidelines for the Methodology and Interpretation of the 6 Minute Walk Test (6MWT). *Adv Respir Med*. 2015;83(4):283–297. doi:10.5603/PiAP.2015.0048
33. He N, Zhang Y, Zhang L, Zhang S, Ye H. Relationship between sarcopenia and cardiovascular diseases in the elderly: An overview. *Front Cardiovasc Med*. 2021; 8:743710. doi:10.3389/fcvm.2021.743710
34. Miarka M, Gibiński K, Janik MK, et al. Sarcopenia: The impact on physical capacity of liver transplant patients. *Life (Basel)*. 2021;11(8):740. doi:10.3390/life11080740
35. Singer JP, Diamond JM, Anderson MR, et al. Frailty phenotypes and mortality after lung transplantation: A prospective cohort study. *Am J Transplant*. 2018;18(8):1995–2004. doi:10.1111/ajt.14873
36. Gillis C, Ljungqvist O, Carli F. Prehabilitation, enhanced recovery after surgery, or both? A narrative review. *Br J Anaesth*. 2022;128(3): 434–448. doi: 10.1016/j.bja.2021.12.007
37. Dourakis SP, Geladari E, Geladari C, Vallianou N. Cirrhotic cardiomyopathy: The interplay between liver and cardiac muscle. How does the cardiovascular system react when the liver is diseased? *Curr Cardiol Rev*. 2021;17(1):78–84. doi:10.2174/1573403X15666190509084519
38. Duong N, Sadowski B, Rangnekar AS. The impact of frailty, sarcopenia, and malnutrition on liver transplant outcomes. *Clin Liver Dis*. 2021;17(4):271–276. doi:10.1002/cld.1043
39. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the journal of cachexia, sarcopenia and muscle: Update 2017. *J Cachexia Sarcopenia Muscle*. 2017;8(6):1081–1083. doi:10.1002/jcsm.12261
40. King J, Lowery DR. Physiology, cardiac output. In: *StatPearls*. Treasure Island, USA: StatPearls Publishing; 2025: Bookshelf ID: NBK470455. http://www.ncbi.nlm.nih.gov/books/NBK470455. Accessed October 13, 2025.
41. Kobe J, Mishra N, Arya V, Al-Moustadi W, Nates W, Kumar B. Cardiac output monitoring: Technology and choice. *Ann Card Anaesth*. 2019; 22(1):6. doi: 10.4103/aca.ACA_41_18
42. Sangkum L, Liu GL, Yu L, Yan H, Kaye AD, Liu H. Minimally invasive or noninvasive cardiac output measurement: An update. *J Anesth*. 2016;30(3):461–480. doi:10.1007/s00540-016-2154-9
43. Glass C, Hipskind P, Tsien C, et al. Sarcopenia and a physiologically low respiratory quotient in patients with cirrhosis: A prospective controlled study. *J Appl Physiol*. 2013;114(5):559–565. doi:10.1152/japplphysiol.01042.2012
44. Dasarathy S, Merli M. Sarcopenia from mechanism to diagnosis and treatment in liver disease. *J Hepatol*. 2016;65(6):1232–1244. doi: 10.1016/j.jhep.2016.07.040
45. Ebadi M, Bhanji RA, Mazurak VC, Montano-Loza AJ. Sarcopenia in cirrhosis: From pathogenesis to interventions. *J Gastroenterol*. 2019; 54(10):845–859. doi:10.1007/s00535-019-01605-6

Oncogenic impact of *PIK3CA*, *KRAS*, and *PTEN* mutations in cervical cancer among South Indian women

Akram Husain Rehman Syed Rasheed[1,A–D], Praveen Kumar Chandra Sekar[1,A–D,E], Anoop Sreevalsan[2,B], Ramakrishnan Veerabathiran[1,C,F], Vijaya Anand[3,C]

[1] Human Cytogenetics and Genomics Laboratory, Faculty of Allied Health Sciences, Chettinad Hospital and Research Institute, Chettinad Academy of Research and Education, Kelambakkam, India
[2] Department of Obstetrics and Gynecology, Chettinad Hospital and Research Institute, Chettinad Academy of Research and Education, Kelambakkam, India
[3] Department of Human Genetics and Molecular Biology, Bharathiar University, Coimbatore, India

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):997–1008

Address for correspondence
Ramakrishnan Veerabathiran
E-mail: rkgenes@gmail.com

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
The authors thank all study participants for their participation in this study. The authors also thank the Chettinad Academy of Research and Education for its continued support and encouragement.

Received on January 21, 2025
Reviewed on July 16, 2025
Accepted on August 19, 2025

Published online on June 5, 2026

Cite as
Rehman Syed Rasheed AH, Chandra Sekar PK, Sreevalsan A, Veerabathiran R, Anand V. Oncogenic impact of *PIK3CA*, *KRAS*, and *PTEN* mutations in cervical cancer among South Indian women. *Adv Clin Exp Med*. 2026;35(6):997–1008. doi:10.17219/acem/209713

DOI
10.17219/acem/209713

Abstract

Background. Cervical cancer (CC) affects millions of women worldwide. This condition is strongly associated with human papillomavirus (HPV) infection. Oncogenic alterations are known to contribute to the development and progression of CC.

Objectives. This study aimed to screen for variations in selected genes associated with CC, including *PIK3CA*, *KRAS*, and *PTEN*, and to detect high-risk HPV genotypes 16, 18, 31, 45, 52, and 58 using gene-specific polymerase chain reaction (PCR), followed by single-strand conformation polymorphism (SSCP) analysis and confirmation using bidirectional DNA sequencing.

Materials and methods. The study included 414 participants, comprising 204 cases and 210 controls. Healthy controls were disease-free individuals participating in regular health checkups. Selected gene mutations were analyzed using PCR-based assays, SSCP, and sequence analysis. HPV genotyping was also performed.

Results. All study participants were analyzed for mutations in the *PIK3CA*, *KRAS*, and *PTEN* genes. The analysis revealed mutation frequencies of 6.37% for *KRAS*, 2.45% for *PTEN*, and 16.66% for *PIK3CA* in CC cases. These findings suggest that *PIK3CA* and *KRAS* mutations are more frequent in patients with CC than *PTEN* mutations. HPV infection was detected in 87.10% of patients with CC, 79.24% of participants with high-grade squamous intraepithelial lesions (HSIL), and 60.34% of participants with low-grade squamous intraepithelial lesions (LSIL). This study contributes to understanding the genetic basis of CC in South India and may facilitate the development of future targeted therapies.

Conclusions. The high prevalence of HPV underscores its etiological significance in CC. These findings contribute to a deeper understanding of the molecular mechanisms underlying CC in this population and may support the development of targeted therapeutic strategies for high-risk individuals. Future prospective studies and functional analyses are warranted to validate the clinical significance of these mutations and clarify their role in disease progression.

Keywords: genetics, cervical cancer, oncogenes, genotyping, mutations

Highlights

- This study investigates genetic mutations in the *PIK3CA*, *KRAS*, and *PTEN* genes in cervical cancer patients from South India.
- A higher prevalence of *PIK3CA* mutations (16.66%) was observed in cervical cancer cases compared with *KRAS* (6.37%) and *PTEN* (2.45%) mutations.
- High-risk HPV 16 and 18 genotypes were associated with genetic mutations and histopathological grades of cervical cancer.
- Mutation profiling using PCR, single-stranded conformation polymorphism (SSCP), and sequencing revealed potential biomarkers for personalized therapies.
- The results highlight the need for genetic testing to improve cervical cancer diagnosis and management.

Background

Cervical cancer (CC) is a major gynecological malignancy that arises in the lower part of the uterus, connecting the corpus uteri to the vagina.[1] Despite advances in screening and vaccination programs, CC remains a significant public health burden, particularly in low- and middle-income countries.[2] By 2026, an estimated 528,000 new cases of CC are expected, with developing nations contributing approx. 85% of the global burden.[3] The disease accounts for about 266,000 deaths annually, representing 8% of all cancer-related mortality in women.[4] In the USA alone, approx. 11,500 new cases are diagnosed each year, leading to 4,000 deaths.[5] While CC incidence has declined in many developed countries due to effective screening programs,[6] it remains the 4th most common malignancy among women worldwide, with 604,000 new cases and 342,000 deaths reported in 2020.[7] India and China together account for over 1/3 of the global CC burden, with India reporting 97,000 new cases and 60,000 deaths annually. Notably, India has the highest age-standardized incidence rate in South Asia (22 per 100,000), with regional variations observed between South India (16.7–18.9 per 100,000) and the northeastern part of this country (24.3 per 100,000).[8]

Persistent infection with high-risk human papillomavirus (HPV) types is the primary cause of CC, contributing to more than 90% of cases.[9] Among these, HPV 16 is the most oncogenic, followed by HPV 18.[10] However, like many other malignancies, CC is not solely driven by viral infection but also by genetic alterations that dysregulate key signaling pathways. Mutations in oncogenes and tumor suppressor genes play a crucial role in tumor initiation and progression by disrupting cellular processes such as proliferation, metabolism, and apoptosis.[11] Previous studies have identified genetic variations in multiple cancer-related genes, including *AKT1*, *KRAS*, *HRAS*, *NRAS*, *PIK3CA*, *FGFR2*, *FGFR3*, *HER2*, *BRAF*, *EGFR*, *CCDC6*, and *PTEN*, in patients with CC in China.[12] However, research on the mutational profiles of *PIK3CA*, *KRAS*, and *PTEN* in South Indian patients with CC remains limited, leaving a gap in our understanding of the genetic landscape of this disease in this population.[13]

The *PIK3CA* gene, a key component of the phosphoinositide 3-kinase (PI3K) signaling pathway, is frequently mutated in various cancers, making it a potential biomarker for targeted therapies. The *KRAS* oncogene plays a critical role in tumorigenesis by regulating cell growth and proliferation, and its mutational status has clinical significance in multiple cancer types.[14] The *PTEN* gene, a well-established tumor suppressor, regulates cell migration and inhibits uncontrolled tumor growth.[15] Understanding the mutational status of these genes in CC could provide valuable insights into tumor biology and facilitate the development of more effective therapeutic strategies.

In regions with low socioeconomic status, the incidence and mortality of CC remain disproportionately high due to limited access to early diagnosis and personalized treatment.[16] The limited availability of genetic profiling in South Indian patients with CC may complicate treatment decision-making and restrict opportunities for more individualized therapeutic approaches.[17]

Objectives

This study hypothesizes that mutations in *PIK3CA*, *KRAS*, and *PTEN* contribute to CC pathogenesis in South Indian women. The primary objective was to screen for genetic alterations in *PIK3CA*, *KRAS*, and *PTEN*, alongside high-risk HPV genotyping (types 16, 18, 31, 45, 52, and 58), and to assess their potential role in disease initiation and progression. By identifying specific mutations, this study aims to enhance understanding of the molecular mechanisms underlying CC in an understudied population and provide a foundation for future targeted therapeutic strategies.

Materials and methods

Sample collection

Samples for the study were collected from patients attending the Department of Obstetrics and Gynecology

at Chettinad Hospital and Research Institute and Chettinad Super Speciality Hospital (CSSH) in Coimbatore, India. Patients were screened for HPV infection using type-specific polymerase chain reaction (PCR). The study protocol was approved by the Institutional Human Ethics Committee of the Chettinad Academy of Research and Education (CARE-IHEC; approval No. IHEC/D.NO:012), and informed consent was obtained from all participants prior to sample collection.

Inclusion and exclusion criteria

The study included patients diagnosed with CC across various histopathological grades, as well as healthy controls with no prior history of CC, HPV infection, or other gynecological disorders. Participants with a history of conization or hysterectomy, those who were pregnant, and individuals with severe comorbid conditions, sexually transmitted diseases, or other malignancies were excluded.

Sampling

Cervical scrapings were collected from study participants using a sterile disposable cervical cytobrush into a clean collection container (Citotest Labware Manufacturing Ltd., Haimen, China) for HPV detection and genetic analysis. Participants were also examined by trained gynecologists at Chettinad Hospital and Research Institute (Kelambakkam, India), who performed colposcopic examinations and obtained tissue samples when abnormalities were identified. The sample size was determined through power analysis using the formula $n = p(1 - p)(Z/E)^2$, with a statistical power of 0.95.[18] A total of 414 individuals were included, comprising 204 cases and 210 controls. The control group consisted of healthy individuals undergoing routine health checkups who were free of disease.

Extraction of genomic DNA

Genomic DNA was extracted from cervical scrapings using the phenol-chloroform method.[19] DNA quantity and quality were assessed using spectrophotometric analysis, and gel electrophoresis was performed for further evaluation.

Mutation profiling and Sanger sequencing for mutation confirmation

According to the COSMIC database, previous studies have identified somatic mutations in *PIK3CA*, *PTEN*, and *KRAS* in cervical carcinoma.[20] Therefore, these 3 genes were selected as targets for the present study, and gene-specific PCR primers were used to amplify mutation hotspot regions. Mutation detection within oncogenic hotspot regions was performed using gene-specific PCR.

Subsequent analyses included single-strand conformation polymorphism (SSCP) analysis and confirmation with bidirectional DNA sequencing. The nucleotide sequences of the exonic regions of *PIK3CA*, *KRAS*, and *PTEN* were retrieved from the Ensembl genome database (https://www.ensembl.org/index.html).[21] Primers were newly designed for this study using the Primer3 (v. 4.0) program (https://primer3.ut.ee). The quality of the designed primers was evaluated using online tools, including Primer Stats and OligoCalc (https://www.bioinformatics.org/sms2/pcr_primer_stats.html and https://www.biosyn.com/gizmo/tools/oligo/oligonucleotide%20properties%20calculator.htm). The designed primer sequences and their properties are presented in Table 1. Polymerase chain reaction products were separated on a 1.2% agarose gel to visualize and analyze the specific amplicon bands. The amplified fragments were excised from the agarose gel and purified using the QIAquick PCR Purification Kit (cat. No. 28104) (Qiagen, Hilden, Germany). After purification, the DNA fragments were sequenced using the Sanger sequencing method[22] with the Applied Biosystems 3130 system (Applied Biosystems, Foster City, USA). The obtained sequences were then compared with the reference sequences from the Ensembl genome database to identify mutations.

Statistical analyses

All statistical analyses were performed using IBM SPSS v. 21 (IBM Corp., Armonk, USA), with statistical significance set at $p < 0.05$. The normality of continuous variables, such as age, was assessed using the Shapiro–Wilk test, and appropriate parametric tests (Student's t-test) were applied where assumptions of normality were met. Categorical variables were compared using Pearson's χ^2 test, while Fisher's exact test was applied when expected cell counts were <5. For ordinal variables, trend analysis was conducted using the Cochran–Mantel–Haenszel (CMH) test.

Univariate risk ratios (RRs) with 95% confidence intervals (95% CIs) were calculated for selected demographic and lifestyle variables (e.g., tobacco use, contraceptive use, parity, and family history) to assess their association with CC risk. Multivariable logistic regression models with elastic net regularization (combining L1 and L2 penalties) were used to examine associations between mutations in *PIK3CA*, *KRAS*, and *PTEN* and clinical/lifestyle predictors, including HPV status, age, parity, contraceptive use, tobacco use, and family history of cancer. Ten-fold cross-validation was used to optimize the penalty parameters. Final model results are reported as adjusted odds ratios (aORs), 95% CIs, and p-values. Model classification performance was assessed using precision, recall, and F1 score at a probability threshold of 0.3. Receiver operating characteristic (ROC) curves and areas under the curve (AUCs) are presented in the supplementary materials.

Table 1. Designed primer sequences for selected oncogenes

Sample No.	Name	Sequences of PCR primers (5′-3′)	Total number of bases	GC content (%)	Size of the PCR product (bp)
1.	*PIK3CA*-Exon 1-F	GAACTGTGGGGCATCCACTT	20	55	299
2.	*PIK3CA*-Exon 1-R	CGGTTGCCTACTGGTTCAAT	20	50	
3.	*PIK3CA*-Exon 9-F	CTGTGAATCCAGAGGGGAAA	20	50	202
4.	*PIK3CA*-Exon 9-R	TTAGCACTTACCTGTGACTC	20	45	
5.	*PIK3CA*-Exon 20-F	GCTCCAAACTGACCAAACTG	20	20	382
6.	*PIK3CA*-Exon 20-R	GGTCTTTGCCTGCTGAGAGT	20	55	
7.	*KRAS*-Exon 2-F	CTGGTGGAGTATTTGATAGTG	21	43	203
8.	*KRAS*-Exon 2-R	CTCTATTGTTGGATCATATTCGTC	24	38	
9.	*KRAS*-Exon 3-F	TGTGTTTCTCCCTTCTCAGGA	21	48	275
10.	*KRAS*-Exon 3-R	GGCATTAGCAAAGACTCA	18	44	
11.	*KRAS*-Exon 4-F	TTGTGGACAGGTTTTGAAAGA	21	38	237
12.	*KRAS*-Exon 4-R	GTTACTTACCTGTCTTGTCTTTGC	24	42	
13.	*PTEN*-Exon 5-F	AGACCATAACCCACCACAGC	20	55	209
14.	*PTEN*-Exon 5-R	TGGTCCTTACTTCCCCATAGAA	22	45	
15.	*PTEN*-Exon 9-F	TAGAGGAGCCGTCAAATCCA	20	50	195
16.	*PTEN*-Exon 9-R	TCATGGTGTTTTATCCCTCTTG	22	41	

F – forward primer; R – reverse primer; bp – base pair; PCR – polymerase chain reaction; GC – guanine and cytosine.

Results

Clinical features of the cases and controls

The demographic and baseline characteristics of the study groups are presented as mean ± standard deviation (SD) in Table 2. The mean age of the participants was approximately normally distributed, as indicated by the Shapiro–Wilk test ($p = 0.135$ for cases; $p = 0.838$ for controls). As age was approximately normally distributed and differed modestly between groups, a 2-sample Student's t-test was used for comparison. The analysis showed a significant difference between the case group (mean age: 53 ±8.90 years) and the control group (mean age: 50 ±5.49 years) ($t = 4.14$, degrees of freedom (df) = 412, $p < 0.001$).

Among the categorical variables, having more than 4 pregnancies ($\chi^2 = 11.19$, df = 1, $p = 0.001$), contraceptive use ($\chi^2 = 9.93$, df = 1, $p = 0.002$), family history of cancer ($\chi^2 = 5.70$, df = 1, $p = 0.017$), and tobacco use ($\chi^2 = 7.38$, df = 1, $p = 0.007$) were significantly more common in cases than in controls.

No significant differences between cases and controls were detected for education level ($\chi^2 = 1.65$, df = 3, $p = 0.65$), employment status ($\chi^2 = 1.79$, df = 2, $p = 0.41$), socioeconomic status ($\chi^2 = 2.01$, df = 2, $p = 0.37$), or the composite comorbidity category ($\chi^2 = 2.28$, df = 3, $p = 0.52$). In addition, risk ratios (RRs) with 95% CIs were calculated for binary variables. Having more than 4 pregnancies (RR = 2.14), contraceptive use (RR = 2.65), and tobacco use (RR = 2.27) showed strong positive associations with CC risk. Group differences in clinical and lifestyle factors were assessed using χ^2 tests, and variables showing significant associations were considered potential confounders in the regression analysis. The normality of continuous variables, including age, was assessed using the Shapiro–Wilk test. The results are presented in Supplementary Table 1. All p-values in this section were derived from univariate analyses using Pearson's χ^2 test, where appropriate.

Genotyping of 6 high-risk HPVs

HPV genotyping was performed under optimized PCR conditions. Amplicons ranging from 150 to 295 base pairs indicated the presence of specific HPV genotypes in both cases and controls. Polymerase chain reaction products were analyzed using a 100 bp DNA ladder on a 1.2% agarose gel (Fig. 1). The PCR findings were subsequently correlated with histopathological grading.

Pearson's χ^2 test revealed a highly significant association between HPV infection and histopathological grade ($\chi^2 = 108.56$, df = 3, $p < 0.001$), with HPV positivity increasing from negative for intraepithelial lesion or malignancy (NILM; 10.47%) to low-grade squamous intraepithelial lesion (LSIL; 60.34%), high-grade squamous intraepithelial lesion (HSIL; 79.24%), and CC (87.10%) (Table 3). A CMH test confirmed a significant increasing trend in HPV positivity with worsening histopathological grade (p for trend < 0.001). Among HPV genotypes, HPV 16 and HPV 58 showed significant variation across histopathological grades (HPV 16: $\chi^2 = 8.10$, df = 2, $p = 0.017$; HPV 58: $\chi^2 = 9.11$, df = 2, $p = 0.033$), with higher frequencies in more advanced lesions. No statistically significant distribution patterns were observed for HPV 18, 31, 45, or 52 ($p > 0.05$ for all) (Table 4).

Table 2. Demographic characteristics of the study subjects

Sample No.	Factors	Cases (n = 204)	Controls (n = 210)	Test (statistic, df)	p-value*	RR	95% CI
1.	mean age	53 ±8.90[a]	50 ±5.49[a]	Student's t-test = 4.14, df = 412	p < 0.001	–	–
2.	>4 pregnancies	50 (24.50)[b]	24 (11.4)[b]	χ^2 = 11.19, df = 1	p < 0.001	2.14	(1.41–4.19)
3.	usage of contraceptives	35 (17.15)[b]	14 (6.6)[b]	χ^2 = 9.93, df = 1	p = 0.002	2.65	(1.33–5.04)
4.	family history of any cancer	18 (8.82)[b]	6 (2.9)[b]	χ^2 = 5.70, df = 1	p = 0.017	3.04	(0.80–2.58)
5.	tobacco usage	33 (16.17)[b]	15 (7.1)[b]	χ^2 = 7.38, df = 1	p = 0.007	2.27	(1.51–5.01)
6.	level of education						
	primary	24 (11.76)[b]	17 (8.09)[b]	χ^2 = 1.65, df = 3	p = 0.648	–	–
	high school	67 (32.84)[b]	69 (32.8)[b]				
	higher secondary	64 (31.37)[b]	70 (33.3)[b]				
	graduate	49 (24.01)[b]	54 (25.7)[b]				
7.	work status						
	employed	38 (18.62)[b]	43 (20.4)[b]	χ^2 = 1.79, df = 2	p = 0.409	–	–
	retired	29 (14.21)[b]	21 (10.4)[b]				
	homemaker	137 (67.15)[b]	146 (69.4)[b]				
8.	socioeconomic status						
	slow	84 (41.17)[b]	6 (36.2)[b]	χ^2 = 2.01, df = 2	0.367		
	middle	103 (50.49)[b]	109 (51.9)[b]				
	high	17 (8.33)[b]	25 (11.9)[b]				
9.	other complications						
	hypertension	08 (3.92)[b]	05 (2.3)[b]	χ^2 = 2.28, df = 3	p = 0.516	–	–
	T2DM	10 (4.90)[b]	03 (1.4)[b]				
	thyroid	14 (6.86)[b]	11 (5.2)[b]				
	obesity	20 (9.80)[b]	08 (3.8)[b]				

[a]data are presented as mean ± standard deviation (SD); [b]percentage; df – degrees of freedom; RR – risk ratio; 95% CI – 95% confidence interval; T2DM – type 2 diabetes mellitus;*p-values were derived using Pearson's χ^2 test of independence. Only applicable binary variables are presented with RR and 95% CI. Values in bold are statistically significant.

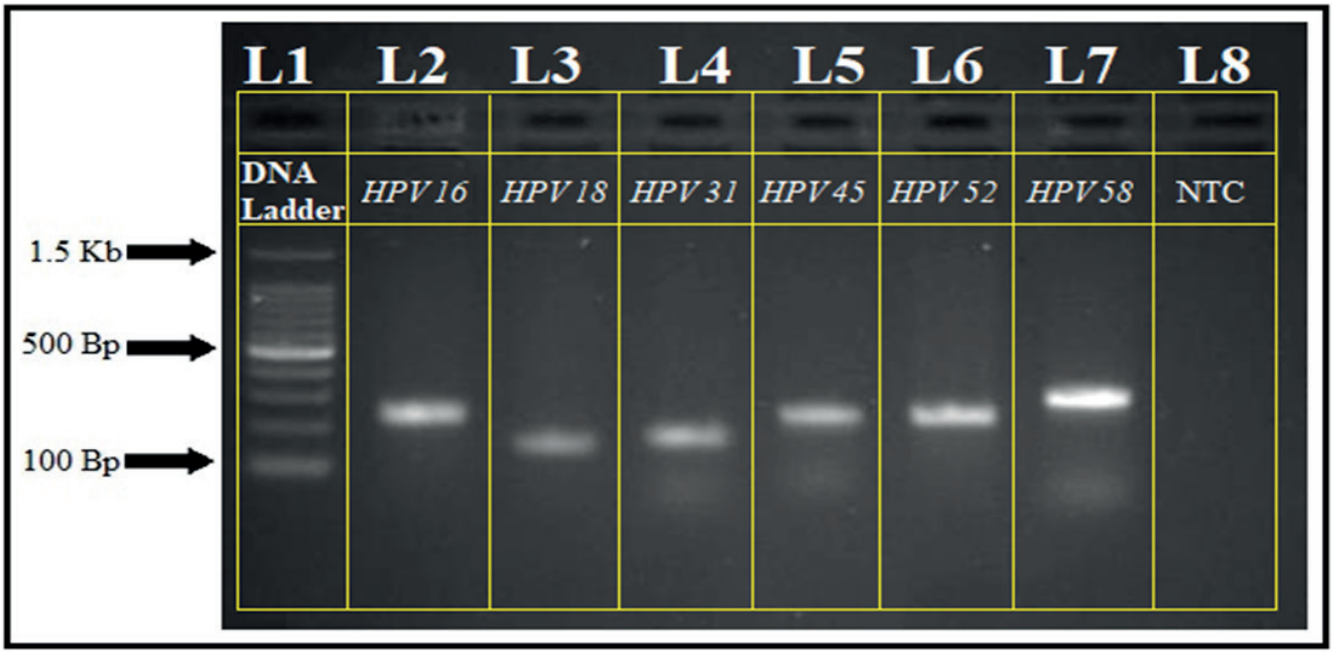


Fig. 1. Optimized polymerase chain reaction (PCR) conditions for high-risk human papillomavirus (HPV) genotypes

Table 3. HPV identification by type-specific polymerase chain reaction (PCR)

Histopathological grade	Samples tested	HPV-positive	Percentage (%)	HPV-negative	p-value
Negative for intraepithelial lesion or malignancy (NILM)	210	22	10.47	188	–
Low-grade squamous intraepithelial lesion (LSIL)	58	35	60.34	23	<0.001
High-grade squamous intraepithelial lesion (HSIL)	53	42	79.24	11	–
Cervical cancer (CC)	93	81	87.10	12	–

*p-value derived from Cochran–Mantel–Haenszel (CMH) test ($\chi^2 = 108.56$, df = 3; $p < 0.001$); HPV – human papillomavirus; df – degrees of freedom.

Table 4. HPV genotype distribution among the LSIL, HSIL, and CC

HPV genotypes	LSIL, % (n)	HSIL, % (n)	CC, % (n)	χ^2, df	p-value
16	40.00 (14)	42.80 (23)	67.90 (55)	$\chi^2 = 8.10$, df = 2	**0.034**
18	25.71 (9)	14.30 (10)	19.75 (16)	$\chi^2 = 0.59$ df = 2	0.786
31	5.71 (2)	4.76 (2)	–	$\chi^2 = 4.38$, df = 2	0.168
45	–	7.14 (3)	7.40 (6)	$\chi^2 = 2.71$, df = 2	0.308
52	11.42 (4)	4.76 (2)	2.46 (2)	$\chi^2 = 4.09$, df = 2	0.308
58	17.17 (6)	4.76 (2)	2.46 (2)	$\chi^2 = 9.11$, df = 2	**0.033**

*p-values derived from Pearson's χ^2 test of independence; p-values were adjusted using the Benjamini–Hochberg false discovery rate (FDR) correction for multiple comparisons. NILM – negative for intra-epithelial lesion (NILM); CC – cervical cancer; df – degrees of freedom; HPV – human papillomavirus. Values in bold are statistically significant.

Table 5. Elastic net logistic regression for *PIK3CA* mutations adjusted for confounding factors

Predictors	aOR (*PIK3CA*)	95% CI	p-value	aOR (*KRAS*)	95% CI	p-value
Cervical cancer	74.03	6.39–857.89	<0.001	7.44	1.41–39.36	**0.018**
HPV-positive	1.20	0.60–2.38	0.593	0.79	0.44–1.42	0.432
>4 pregnancies	0.84	0.51–1.37	0.488	0.26	0.17–0.40	<0.001
Contraceptive use	0.89	0.52–1.51	0.664	0.76	0.43–1.34	0.340
Family history	0.14	0.05–0.37	<0.001	0.48	0.20–1.16	**0.103**
Tobacco use	1.11	0.62–2.00	0.725	1.34	0.73–2.46	0.342

aOR – adjusted odds ratio. 95% CI – 95% confidence interval; HPV – human papillomavirus. Elastic net regularization (L1 + L2) was applied to logistic regression models to control for overfitting and sparse data; p-values derived from the logistic model. Bold values indicate stronger associations.

Multivariable analysis of gene mutations

To assess the independent contribution of clinical and lifestyle factors to mutation risk, elastic net-regularized logistic regression was applied for *PIK3CA* and *KRAS* mutations. This method combines L1 and L2 penalties to balance variable selection and coefficient shrinkage and is particularly suitable for small or sparse datasets. The model identified CC status as the strongest independent predictor of both *PIK3CA* (aOR = 74.03) and *KRAS* (aOR = 7.44) mutations. Other variables, including HPV positivity, tobacco use, parity, contraceptive use, and family history of cancer, were retained in the models but did not demonstrate statistically significant associations (Table 5).

Mutation profiling

Mutations in the *PIK3CA*, *KRAS*, and *PTEN* genes were analyzed among study participants using gene-specific PCR. The amplicon sizes ranged from 180 to 400 base pairs, and the PCR products were confirmed with agarose gel electrophoresis (Fig. 2). Mutation frequency comparisons between cases and controls were assessed using univariate ORs.

The observed mutation frequencies among CC cases were 16.66% for *PIK3CA* (95% CI: 12.34–20.98, p = 0.002), 6.37% for KRAS (95% CI: 4.11–8.63, p = 0.017), and 2.45% for PTEN (95% CI: 1.12–3.78, p = 0.042). No significant mutations were identified in the control group, resulting in perfect separation for some outcomes. Statistical comparisons between cases and controls confirmed a strong association between *PIK3CA* mutations and CC risk (OR = 3.89, 95% CI: 2.14–7.08, p < 0.001).

Gene profiling of *PIK3CA*, *KRAS*, and *PTEN*

Multivariable logistic regression was performed to assess the association between gene mutations and potential confounding factors. Penalized (L2) logistic regression was applied due to sparse mutation events and perfect separation in some models. Results are presented as aORs with 95% CIs. The distribution of mutations in *PIK3CA*, *KRAS*, and *PTEN* was further analyzed in relation to HPV genotype

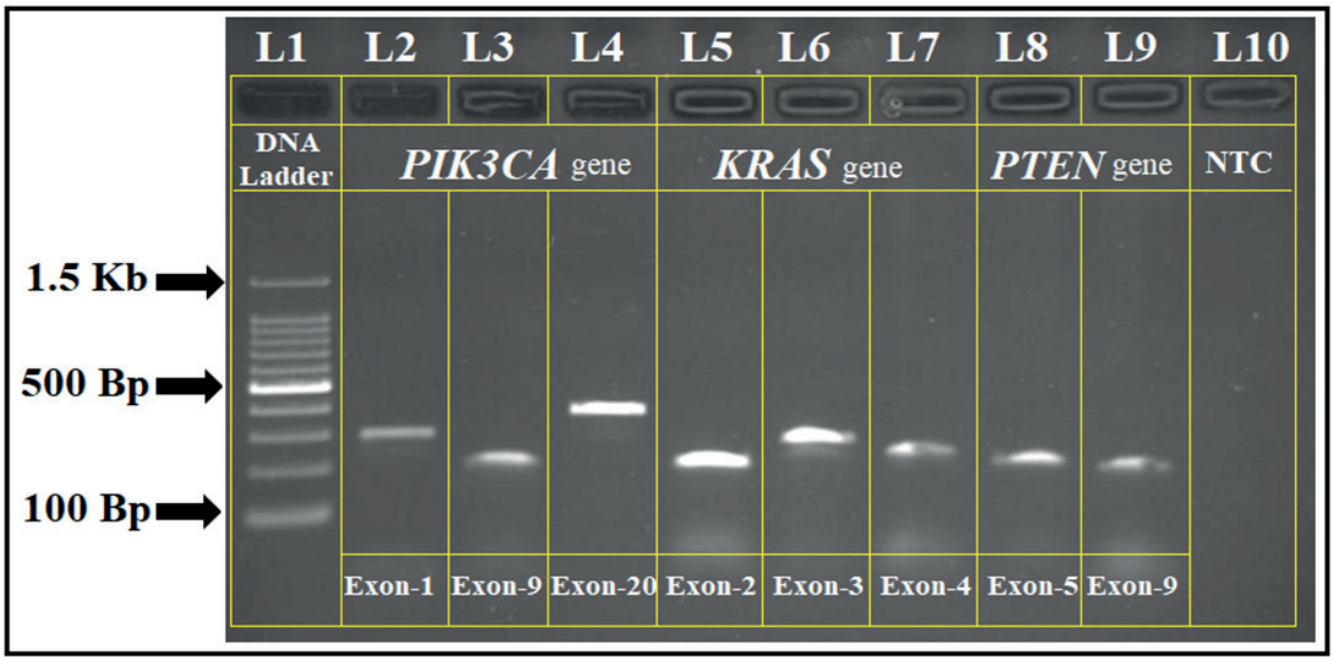


Fig. 2. Optimized polymerase chain reaction (PCR) conditions for the selected genes analyzed in this study

and histopathological grade (Fig. 3). *PIK3CA* mutations were predominantly identified in exons 9, 20, and 1, with hotspot mutations located in the helical domain (*E542K*, *E545K*) and catalytic domain (*H1047R*). The overall frequency of these mutations was 2.94%. Patients harboring *PIK3CA* mutations had a significantly higher likelihood of HPV 16 infection ($p = 0.004$). Detailed information regarding *PIK3CA* mutations in relation to HPV genotypes and histopathological grading is presented in Table 6. *KRAS* mutations were primarily detected in exon 2, followed by exons 3 and 4 (Table 7). These mutations were significantly associated with a family history of breast cancer and consanguineous marriage ($p = 0.016$). As shown in Table 8, logistic regression analysis demonstrated that *KRAS* mutations were significantly associated with HPV positivity (OR = 2.47, 95% CI: 1.32–4.21, $p = 0.022$) and CC (OR = 3.80, 95% CI: 2.01–6.88, $p < 0.001$). An interaction term between HPV status and cancer status was also statistically significant.

PTEN gene variants were predominantly identified in exon 5, followed by exon 9. These mutations were detected exclusively in CC cases and were not observed in any control subjects. Patients harboring *PTEN* mutations did not report a family history of cancer or consanguinity, suggesting that these variants may represent somatic alterations contributing to carcinogenesis. A novel *PTEN* mutation identified in this study has not previously been reported in HPV-associated CC (Table 9).

Model performance evaluation

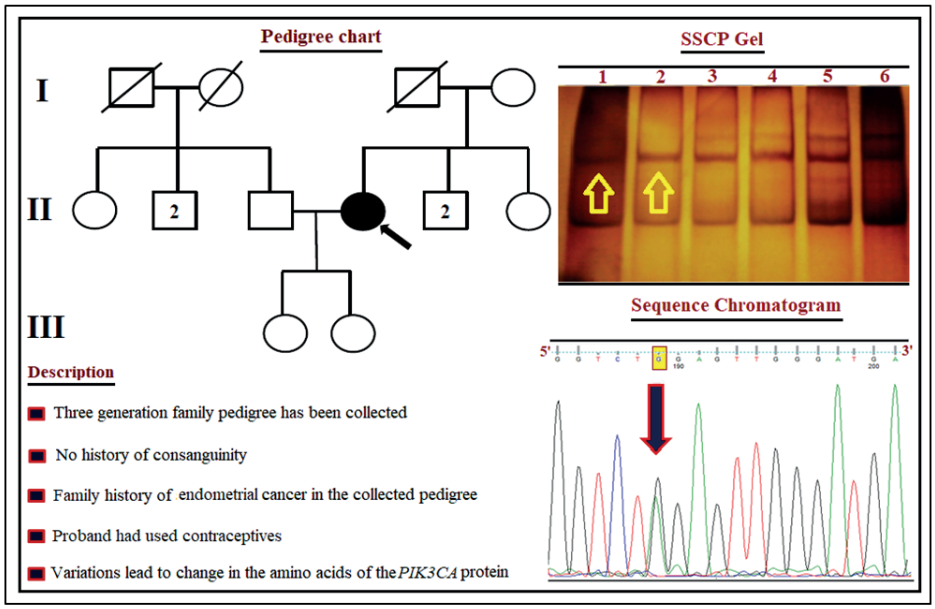
To further evaluate the classification performance of the mutation prediction models, various probability thresholds (e.g., 0.3) were tested. However, due to extreme class imbalance, the models failed to identify mutation-positive cases (zero true positives). Despite this limitation, overall accuracy remained relatively high (91.8% for *PIK3CA* and 96.9% for *KRAS*), largely driven by the high number of correctly classified negative cases. These results are presented in Supplementary Table 2. Nevertheless, ROC analysis indicated that both models retained moderate-to-strong discriminatory ability, with AUC values of 0.825 for *PIK3CA* and 0.881 for *KRAS* (Supplementary Table 3). The corresponding ROC curves are shown in Supplementary Fig. 4.

Discussion

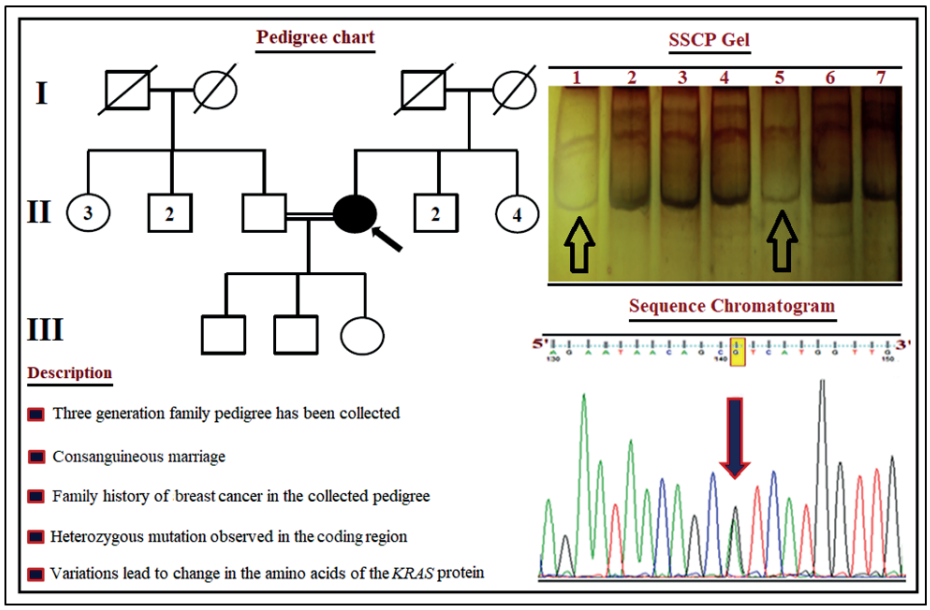
Advances in high-throughput sequencing technologies have enabled the identification of novel genetic variants and somatic mutations with high depth of coverage, contributing to the transition toward precision medicine. The fundamental principle of mutation profiling is the detection of genetic alterations in candidate or disease-associated genes. This approach may facilitate the identification of predictive biomarkers and improve the effectiveness of therapeutic strategies.

The present study aimed to identify and characterize CC-associated mutations using gene-specific PCR, followed by SSCP analysis and confirmation by DNA sequencing. According to our findings, mutations in *PIK3CA* (16.66%), *KRAS* (6.37%), and *PTEN* (2.45%) were associated with CC in South Indian patients. *PIK3CA* is one of the genes frequently implicated in human malignancies. The PI3K signaling pathway plays a critical role in multiple cellular processes, including cell survival, metabolism, growth, and proliferation. *PIK3CA* mutation

A. *PIK3CA*



B. *KRAS*



C. *PTEN*

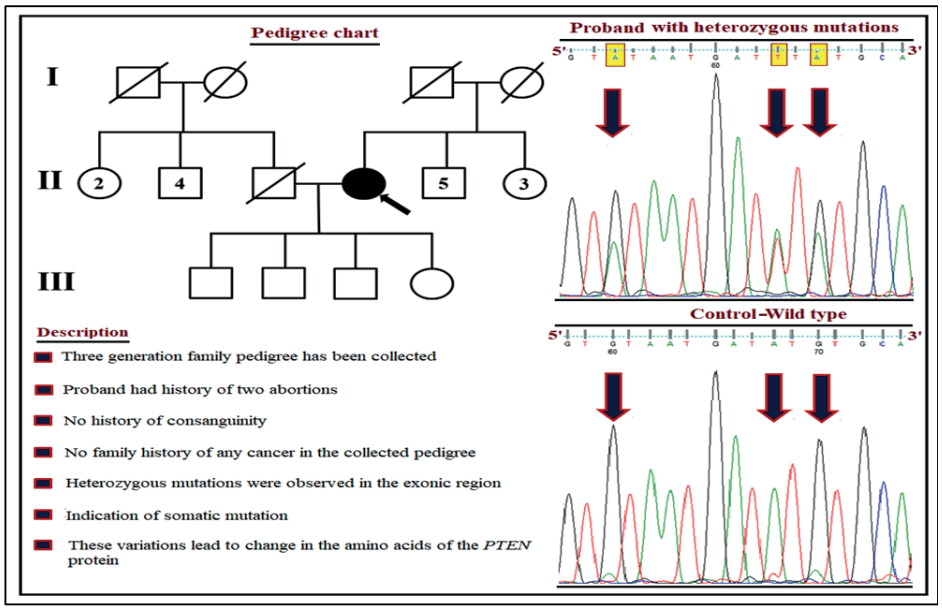


Fig. 3. A. Genetic analysis of *PIK3CA* mutations in cervical cancer (CC); B. Genetic analysis of *KRAS* mutations in CC; C. Genetic analysis of *PTEN* mutations in CC

Table 6. Histological characteristics, human papillomavirus (HPV) status, and *PIK3CA* mutations of study subjects

Sample No.	Histological grading	HPV genotype	Exonic region	Nucleotide change	Amino acid change
1.	CC	16	01	*c.35 G > T*	p.G12C
2.	CC	16		*c.70 T > A*	p.C24S
3.	HSIL	16		*c.223 C > G*	p.Q75E
4.	CC	18		*c.100 T > C*	p.L34L
5.	CC	16		*c.241 G > A*	p.E81K
6.	LSIL	18		*c.115 G > A*	p.E39K
7.	CC	16	09	*c.1633 G > A*	p.E545K
8.	CC	18		*c.1624 G > A*	p.E542K
9.	CC	16		*c.1633 G > A*	p.E545K
10.	CC	16		*c.1624 G > A*	p.E542K
11.	HSIL	18		*c.1633 G > A*	p.E545K
12.	CC	45		*c.1624 G > A*	p.E542K
13.	CC	18		*c.1634 A > C*	p.E545A
14.	CC	16		*c.1633 G > A*	p.E545K
15.	HSIL	16		*c.1624 G > A*	p.E542K
16.	LSIL	58		*c.1634 A > C*	p.E545A
17.	CC	16		*c.1624 G > A*	p.E542K
18.	CC	18		*c.1633 G > C*	p.E545Q
19.	HSIL	16		*c.1633 G > A*	p.E545K
20.	CC	16		*c.1634 A > C*	p.E545A
21.	HSIL	18		*c.1624 G > A*	p.E542K
22.	LSIL	16		*c.1633 G > C*	p.E545Q
23.	CC	16		*c.1635 G > T*	p.E545D
24.	CC	45		*c.1633 G > A*	p.E545K
25.	CC	18	20	*c.3140 A > T*	p.H1047L
26.	CC	16		*c.3140 A > G*	p.H1047R
27.	CC	16		*c.3140 A > T*	p.H1047L
28.	LSIL	58		*c.3146 G > C*	p.G1049A
29.	HSIL	18		*c.3140 A > G*	p.H1047R
30.	CC	45		*c.3140 A > T*	p.H1047L
31.	CC	18		*c.3140 A > G*	p.H1047R
32.	LSIL	16		*c.3146 G > C*	p.G1049A
33.	HSIL	16		*c.3140 A > T*	p.H1047L
34.	CC	16		*c.3140 A > G*	p.H1047R

NILM – negative for intra-epithelial lesion (NILM); LSIL – low-grade squamous intra-epithelial lesion; HSIL – high-grade squamous intra-epithelial lesion; CC – cervical cancer.

and amplification are among the most common mechanisms leading to aberrant activation of the PI3K pathway in various cancers.[23] The mutation frequency of *PIK3CA* in our study was 16.66% among patients with CC. When compared with previous studies, a cohort from the Netherlands including 301 patients with CC of Caucasian ancestry reported *PIK3CA* mutations in 20% of cases. Similarly, a study of 213 Chinese patients with CC reported a mutation frequency of 12.3%.[24] Another study involving Chinese patients of Asian ancestry that screened 16 genes also reported a *PIK3CA* mutation frequency of 12.3%.[25] Our observed *PIK3CA* mutation frequency was therefore somewhat higher than that reported in Chinese cohorts, despite both populations being of Asian ancestry. These differences may reflect variation in mutation spectra, patient characteristics, HPV subtype distribution, or methodological differences between studies. A study from the USA involving 67 patients with CC reported *PIK3CA* mutations in 27.1% of cases, which is substantially higher than the frequency observed in our cohort.[26] The higher prevalence of *PIK3CA* mutations in some Western populations may reflect differences in environmental exposures, lifestyle factors, genetic background, or interactions with high-risk HPV variants.

Table 7. Histological characteristics, human papillomavirus (HPV) status, and *KRAS* mutations of study subjects

Sample No.	Histological grading	HPV genotype	Exonic region	Nucleotide change	Amino acid change
1.	CC	16	02	*c.111 G > A*	p.G38D
2.	CC	16		*c.168 C > T*	p.D57C
3.	CC	45		*c.174 A > T*	p.A59C
4.	CC	16	03	*c.299 T > G*	p.I100K
5.	HSIL	16		*c.349 A > T*	p.N117Y
6.	CC	18		*c.310 A > G*	p.I104K
7.	HSIL	16		*c.362 G > T*	p.P121P
8.	HSIL	45	04	*c.453 G>A*	p.G151G
9.	CC	16		*c.470 A>T*	p.L157L
10.	LSIL	16		*c.491G>A*	p.R171S
11.	CC	18		*c.453 G>A*	p.G151G
12.	HSIL	18		*c.480 T>A*	p.C167S
13.	CC	16		*c.491 G>A*	p.R171S

NILM – negative for intra-epithelial lesion (NILM); LSIL – low-grade squamous intra-epithelial lesion; HSIL – high-grade squamous intra-epithelial lesion; CC – cervical cancer.

Table 8. Logistic regression model for predicting *KRAS* mutations based on human papillomavirus (HPV) status and cervical cancer (CC)

Predictor	OR	95% CI	SE	p-value
HPV-positive	2.47	1.32–4.21	0.45	0.022
Cervical cancer	3.80	2.01–6.88	0.49	<0.001
HPV * CC interaction	1.95	1.01–3.45	0.52	0.38

OR – odds ratio; 95% CI – 95% confidence interval; SE – standard error.

The *KRAS* gene is a well-established oncogene, and multiple point mutations have been identified across various malignancies, including breast, cervical, endometrial, liver, and myeloid cancers.[27] The KRAS protein plays a crucial role in regulating the MAPK signaling pathway and may also influence cellular proliferation through interactions with the PI3K–AKT pathway. Our study demonstrated an association between *KRAS* mutations and high-risk HPV 16/18 genotypes. The integration of *KRAS* mutational analysis with HPV genotyping may improve risk stratification and potentially inform prognostic assessment in patients with CC. The mutation frequency of *KRAS* in this study was 6.37% among patients with CC. In comparison, a previous study from Boston analyzing 80 Caucasian patients across 139 cancer-related genes reported an overall mutation rate of 60%, with *KRAS* mutations accounting for 8.8% of cases.[28] In a Chinese study (n = 876), *KRAS* mutations were detected in 3.4% of CC cases,[29] whereas another Chinese study focused on cervical adenocarcinoma reported *KRAS* mutations in 16.6% of patients.[30] These differences may reflect variation in histological subtype, study population, sequencing methodology, or sample composition.

Although *KRAS* mutations have been reported across diverse populations, current evidence remains insufficient to conclude that they represent a conserved oncogenic event in CC irrespective of ethnicity. In addition to mutation-driven signaling pathways, targeted gene-silencing approaches such as *MCT1* inhibition have shown promise in enhancing dendritic cell-mediated immune responses against CC, suggesting potential synergy between immunotherapeutic and molecular targeting strategies.[31] However, our observed association between *KRAS* mutations and HPV 16/18 infection highlights the potential value of integrating viral and genetic profiling in CC risk assessment.

PTEN is a well-established tumor suppressor gene implicated in the development of multiple cancer types and plays an important role in regulating cellular proliferation, survival, and migration. The phosphatase encoded by *PTEN* negatively regulates the PI3K signaling pathway through dephosphorylation of phosphoinositide substrates, thereby

Table 9. Histological characteristics, HPV status, and *PTEN* mutations of study subjects

Sample No.	Histological grading	HPV genotype	Exonic region	Nucleotide change	Amino acid change	Prediction
1.	CC	16	05	*c.397 G > A*	p.V133L	damaging
				c.405 A>T	p.I135I	silent
				c.407 G>A	p.C136Y	damaging
2.	CC	16	05	*c.466 G>T*	p.R160H	pathogenic
3.	HSIL	45	05	*c.1044 A > T*	p.K234B	pathogenic

HPV – human papillomavirus; CC – cervical cancer; HSIL – high-grade squamous intra-epithelial lesion; CC – cervical cancer.

limiting uncontrolled cellular growth and proliferation.[32] In our study, the mutation frequency of *PTEN* was 2.45% among patients with CC. Compared with previous studies, cohorts from Mexico and Norway analyzing 115 CC and control samples reported a *PTEN* mutation frequency of approx. 6%, which is higher than the frequency observed in our cohort.[33] A Japanese study (n = 50) reported *PTEN* mutations in 4.2% of cases,[34] while a study from the USA reported *PTEN* alterations in 8% of CC cases.[35] Another study in Mexican patients (n = 155) identified *PTEN* mutations in 5% of cases.[36]

Observed differences in *PTEN* mutation frequency across studies may reflect variation in study population characteristics, sequencing methodology, histological subtype composition, or sample size rather than ethnicity alone. The reasons for these discrepancies remain uncertain. In addition, the identification of a novel *PTEN* mutation in our study raises the possibility of previously unrecognized genetic alterations in South Indian patients with CC, warranting further investigation.

Limitations of the study

One important limitation of this study is its focus on a specific population from South India, which may limit the generalizability of the findings to other populations or ethnic groups. Additionally, the relatively small sample size, particularly for mutation frequency analyses, may have reduced the statistical power to detect less common genetic variants. Although techniques such as SSCP and bidirectional DNA sequencing were employed, these methods may not detect all potentially relevant genetic alterations, including large structural variants or epigenetic changes. Furthermore, the use of conventional PCR and Sanger sequencing rather than next-generation sequencing (NGS) may have limited the ability to detect rare, novel, or non-hotspot mutations more comprehensively. A major limitation of the mutation prediction models was their inability to accurately identify mutation-positive cases because of the low frequency of mutation events and severe class imbalance. This resulted in zero sensitivity and F1 scores at the evaluated classification thresholds. Therefore, further large-scale studies using more comprehensive genomic approaches are needed to better understand the genetic landscape of CC in South Indian patients and its relationship to other populations.

Conclusions

Our findings underscore the importance of understanding population-specific genetic variation in CC. The higher mutation frequency of *PIK3CA* compared with *KRAS* and *PTEN* suggests that alterations in the PI3K pathway may play a prominent role in cervical carcinogenesis in South Indian women. These results provide valuable insights into the molecular landscape of CC in this population and add to the growing body of evidence suggesting that genetic alterations, together with HPV infection, may influence disease progression and therapeutic response. This study highlights the potential relevance of *PIK3CA*, *KRAS*, and *PTEN* mutations in CC among South Indian women. A higher mutation frequency was observed for *PIK3CA* (16.66%) compared with *KRAS* (6.37%) and *PTEN* (2.45%), suggesting that alterations in the PI3K/AKT/mTOR pathway may play an important role in cervical carcinogenesis in this population. The observed association between *KRAS* mutations and high-risk HPV 16/18 infection warrants further investigation into the biological and potential clinical significance of this finding. Although *PTEN* mutations were less frequent, their established role in tumor suppression suggests potential relevance to disease biology. Integrating genetic profiling with HPV testing may improve molecular characterization and risk stratification in CC. However, the clinical utility of mutation-guided therapeutic strategies, including PI3K-targeted therapies, alternative approaches for *KRAS*-mutant tumors, or the predictive relevance of *PTEN* alterations for immunotherapy response, requires further validation in larger prospective studies. Future research should focus on further investigating these mutations and clarifying their potential clinical relevance. Our findings suggest that more individualized therapeutic strategies for CC warrant exploration. Although current management strategies do not routinely incorporate molecular stratification beyond standard clinicopathological classification, our study suggests that identifying distinct molecular subpopulations within CC may offer opportunities to improve patient stratification and potentially enhance outcomes in both early- and advanced-stage disease.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.16919759. The package contains the following files:

Supplementary Table 1. Normality test results for continuous variables.

Supplementary Table 2. Performance metrics and confusion matrix for elastic net logistic regression models.

Supplementary Table 3. Performance metrics for elastic net logistic regression models.

Data Availability Statement

Data sharing does not apply to this article, as all data are already included in the manuscript.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Akram Husain Rehman Syed Rasheed
https://orcid.org/0009-0004-2561-3601
Praveen Kumar Chandra Sekar https://orcid.org/0009-0008-5346-9597
Ramakrishnan Veerabathiran https://orcid.org/0000-0002-9307-5428
Vijaya Anand https://orcid.org/0000-0001-7485-1586

References

1. Bhatla N, Aoki D, Sharma DN, Sankaranarayanan R. Cancer of the cervix uteri: 2021 update. *Int J Gynecol Obstet*. 2021;155(Suppl 1):28–44. doi:10.1002/ijgo.13865
2. Ganesan S, Subbiah VN, Michael JCJ. Associated factors with cervical pre-malignant lesions among the married fisher women community at Sadras, Tamil Nadu. *Asia Pac J Oncol Nurs*. 2015;2(1):42–50. doi:10.4103/2347-5625.146223
3. Chandra Sekar PK, Thomas SM, Veerabathiran R. The future of cervical cancer prevention: Advances in research and technology. *Explor Med*. 2024;5:384–400. doi:10.37349/emed.2024.00226
4. Ramamoorthy T, Kulothungan V, Sathishkumar K, et al. Burden of cervical cancer in India: Estimates of years of life lost, years lived with disability and disability adjusted life years at national and subnational levels using the National Cancer Registry Programme data. *Reprod Health*. 2024;21(1):111. doi:10.1186/s12978-024-01837-7
5. Division of Cancer Prevention and Control, Centers for Disease Control and Prevention (CDC). Cervical Cancer Statistics. Atlanta, USA: Centers for Disease Control and Prevention (CDC); 2023. https://www.cdc.gov/cancer/cervical/statistics/index.htm. Accessed July 27, 2025.
6. Hull R, Mbele M, Makhafola T, et al. Cervical cancer in low and middle-income countries (Review). *Oncol Lett*. 2020;20(3):2058–2074. doi:10.3892/ol.2020.11754
7. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
8. Nath A, Sathishkumar K, Das P, Sudarshan KL, Mathur P. A clinicoepidemiological profile of lung cancers in India: Results from the National Cancer Registry Programme. *Indian J Med Res*. 2022;155(2):264–272. doi:10.4103/ijmr.ijmr_1364_21
9. Okunade KS. Human papillomavirus and cervical cancer. *J Obstet Gynaecol*. 2020;40(5):602–608. doi:10.1080/01443615.2019.1634030
10. Da Silva RL, Da Silva Batista Z, Bastos GR, et al. Role of HPV 16 variants among cervical carcinoma samples from Northeastern Brazil. *BMC Womens Health*. 2020;20(1):162. doi:10.1186/s12905-020-01035-0
11. Purushothaman P, Uppal T, Verma SC. Human DNA tumor viruses and oncogenesis. In: *Animal Biotechnology*. Amsterdam, the Netherlands: Elsevier; 2020:131–151. doi:10.1016/B978-0-12-811710-1.00007-0
12. Xiang L, Li J, Jiang W, et al. Comprehensive analysis of targetable oncogenic mutations in Chinese cervical cancers. *Oncotarget*. 2015;6(7):4968–4975. doi:10.18632/oncotarget.3212
13. Fruman DA, Chiu H, Hopkins BD, Bagrodia S, Cantley LC, Abraham RT. The PI3K pathway in human disease. *Cell*. 2017;170(4):605–635. doi:10.1016/j.cell.2017.07.029
14. Janakiraman M, Vakiani E, Zeng Z, et al. Genomic and biological characterization of exon 4 KRAS mutations in human cancer. *Cancer Res*. 2010;70(14):5901–5911. doi:10.1158/0008-5472.CAN-10-0192
15. Milella M, Falcone I, Conciatori F, et al. PTEN: Multiple functions in human malignant tumors. *Front Oncol*. 2015;5:24. doi:10.3389/fonc.2015.00024
16. Cetina-Pérez L, Luvián-Morales J, Delgadillo-González M, et al. Sociodemographic characteristics and their association with survival in women with cervical cancer. *BMC Cancer*. 2024;24(1):161. doi:10.1186/s12885-024-11909-3
17. Burmeister CA, Khan SF, Schäfer G, et al. Cervical cancer therapies: Current challenges and future perspectives. *Tumour Virus Res*. 2022;13:200238. doi:10.1016/j.tvr.2022.200238
18. Sobha SP, Kesavarao KE. Progonostic effect of GSTM1/GSTT1 polymorphism in determining cardiovascular diseases risk among type 2 diabetes patients in South Indian population. *Mol Biol Rep*. 2023;50(8):6415–6423. doi:10.1007/s11033-023-08514-1
19. Chu TY, Hwang KS, Yu MH, Lee HS, Lai HC, Liu JY. A research-based tumor tissue bank of gynecologic oncology: Characteristics of nucleic acids extracted from normal and tumor tissues from different sites. *Int J Gynecol Cancer*. 2002;12(2):171–176. doi:10.1046/j.1525-1438.2002.01085.x
20. Sharmin S, Zohura FT, Islam MdS, et al. Mutational profiles of marker genes of cervical carcinoma in Bangladeshi patients. *BMC Cancer*. 2021;21(1):289. doi:10.1186/s12885-021-07906-5
21. Flicek P, Amode MR, Barrell D, et al. Ensembl 2014. *Nucl Acids Res*. 2014;42(D1):D749–D755. doi:10.1093/nar/gkt1196
22. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74(12):5463–5467. doi:10.1073/pnas.74.12.5463
23. Yang J, Nie J, Ma X, Wei Y, Peng Y, Wei X. Targeting PI3K in cancer: Mechanisms and advances in clinical trials. *Mol Cancer*. 2019;18(1):26. doi:10.1186/s12943-019-0954-x
24. Spaans VM, Trietsch MD, Peters AAW, et al. Precise classification of cervical carcinomas combined with somatic mutation profiling contributes to predicting disease outcome. *PLoS One*. 2015;10(7):e0133670. doi:10.1371/journal.pone.0133670
25. Femi OF. Genetic alterations and *PIK3CA* gene mutations and amplifications analysis in cervical cancer by racial groups in the United States. *Int J Health Sci (Qassim)*. 2018;12(1):28–32. PMID:29623014. PMCID:PMC5870313.
26. Voutsadakis IA. *PI3KCA* mutations in uterine cervix carcinoma. *J Clin Med*. 2021;10(2):220. doi:10.3390/jcm10020220
27. Jančík S, Drábek J, Radzioch D, Hajdúch M. Clinical relevance of KRAS in human cancers. *J Biomed Biotechnol*. 2010;2010:150960. doi:10.1155/2010/150960
28. Wright AA, Howitt BE, Myers AP, et al. Oncogenic mutations in cervical cancer: Genomic differences between adenocarcinomas and squamous cell carcinomas of the cervix. *Cancer*. 2013;119(21):3776–3783. doi:10.1002/cncr.28288
29. Jiang W, Xiang L, Pei X, et al. Mutational analysis of *KRAS* and its clinical implications in cervical cancer patients. *J Gynecol Oncol*. 2018;29(1):e4. doi:10.3802/jgo.2018.29.e4
30. Loong HHF, Du N, Cheng C, et al. KRAS G12C mutations in Asia: A landscape analysis of 11,951 Chinese tumor samples. *Transl Lung Cancer Res*. 2020;9(5):1759–1769. doi:10.21037/tlcr-20-455
31. Sui X, Xi X. *MCT1* gene silencing enhances the immune effect of dendritic cells on cervical cancer cells. *Adv Clin Exp Med*. 2024;33(7):739–749. doi:10.17219/acem/171446
32. Guo CY, Xu XF, Wu JY, Liu SF. PCR-SSCP-DNA sequencing method in detecting *PTEN* gene mutation and its significance in human gastric cancer. *World J Gastroenterol*. 2008;14(24):3804. doi:10.3748/wjg.14.3804
33. Ojesina AI, Lichtenstein L, Freeman SS, et al. Landscape of genomic alterations in cervical carcinomas. *Nature*. 2014;506(7488):371–375. doi:10.1038/nature12881
34. Harima Y, Sawada S, Nagata K, Sougawa M, Ostapenko V, Ohnishi T. Mutation of the *PTEN* gene in advanced cervical cancer correlated with tumor progression and poor outcome after radiotherapy. *Int J Oncol*. 2001;18(3). doi:10.3892/ijo.18.3.493
35. Shin J, Kim SH, Yoon J. PTEN downregulation induces apoptosis and cell cycle arrest in uterine cervical cancer cells. *Exp Ther Med*. 2021;22(4):1100. doi:10.3892/etm.2021.10534
36. Lou H, Villagran G, Boland JF, et al. Genome analysis of Latin American cervical cancer: Frequent activation of the PIK3CA pathway. *Clin Cancer Res*. 2015;21(23):5360–5370. doi:10.1158/1078-0432.CCR-14-1837

Association of lipid layer patterns with tear evaporation in smokers, refractive errors, and high body mass index

Meznah S. Almutairi[1,A,C,F], Sarah A. Alghamdi[1,B], Basal H. Altoaimi[1,A,D–F], Martin Rickert[2,C,E,F], Gamal A. El-Hiti[1,A,C–F]

[1] Department of Optometry, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
[2] Indiana University School of Optometry, Bloomington, USA

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1009–1016

Address for correspondence
Gamal A. El-Hiti
E-mail: gelhiti@ksu.edu.sa

Funding sources
This study was supported by the Ongoing Research Funding Program (grant No. ORF-2026-404), King Saud University, Riyadh, Saudi Arabia.

Conflict of interest
None declared

Received on May 22, 2025
Reviewed on May 27, 2025
Accepted on September 2, 2025

Published online on June 10, 2026

Abstract

Background. The stability of the tear film is essential for ocular health. Smoking, refractive errors (RE), and a high body mass index (BMI) are key risk factors for dry eye disease. Dry eye symptoms may be caused by a high tear evaporation rate (TER) and/or a thin lipid layer.

Objectives. To assess the correlation between TER and lipid layer patterns (LLP) within 4 distinct groups: healthy controls, smokers, subjects with RE, and individuals with high BMI.

Materials and methods. The study included 120 subjects aged 18–30 years divided into 4 groups (30 per group; 15 women and 15 men). The ocular surface disease index (OSDI) was assessed, followed by evaluation of LLP using EASYTEAR View+ and TER using the Delfin VapoMeter.

Results. Kruskal–Wallis tests revealed significant between-group differences in OSDI ($\chi^2 = 62.91$, $n = 120$, $p < 0.001$) and LLP ($\chi^2 = 26.59$, $n = 120$, $p < 0.001$), but not in TER ($\chi^2 = 7.20$, $n = 120$, $p = 0.066$). Within-group Kendall tau-b correlations between TER and LLP revealed strong negative associations in the high-BMI ($\tau b = -0.563$, $n = 30$, $p < 0.001$) and smoker groups ($\tau b = -0.457$, $n = 30$, $p = 0.002$). A moderate negative correlation was found in the RE group ($\tau b = -0.287$, $n = 30$, $p = 0.043$), but not in healthy controls ($\tau b = -0.199$, $n = 30$, $p = 0.189$).

Conclusions. The association between TER and LLP was strongly negative in smokers and the high-BMI group, moderate in the RE group, and absent in healthy controls. The inverse relationship between TER and LLP may indicate compromised tear film stability in populations at risk of dry eye disease.

Key words: smokers, tear film, high body mass index, dry eye syndromes, observational correlation study

Cite as
Almutairi MS, Alghamdi SA, Altoaimi BH, Rickert M, El-Hiti GA. Association of lipid layer patterns with tear evaporation in smokers, refractive errors, and high body mass index. *Adv Clin Exp Med*. 2026;35(6):1009–1016. doi:10.17219/acem/210191

DOI
10.17219/acem/210191

Highlights

- The correlation between tear evaporation rate (TER) and lipid layer patterns (LLP) has not been systematically studied.
- A strong negative correlation was found between TER and LLP in smokers and individuals with a high body mass index (BMI).
- A moderate negative correlation was observed between TER and LLP in individuals with refractive errors (RE).
- This research gap highlights the need for further studies to better understand these correlations.

Background

The tear film is crucial for vision and ocular health. Disturbances in tear film stability and osmolarity can lead to several disorders, with dry eye disease being one of the most common, affecting 5–50% of the population.[1,2] Symptoms include foreign body sensation, dryness, irritation, discomfort, and pain.[3] Dry eye disease has significant economic and social impacts, making it important to identify associated factors such as digital screen use, contact lens wear, smoking, systemic illnesses, and certain medications.[4]

Dysfunction of the meibomian glands leads to increased tear evaporation, whereas lacrimal gland dysfunction results in aqueous deficiency.[5,6] Various tests are used to assess dry eye disease, as each measures different parameters.[7–10] The lipid layer spreads across the tear film with each blink to reduce tear evaporation and is therefore crucial for maintaining tear film stability and function.[11] A portable evaporimeter has been effectively used to measure tear evaporation rate (TER).[12,13] This technique is quick, convenient, repeatable, and noninvasive compared with other methods. A TER value exceeding 25 $g/m^2/h$ at room temperature and approx. 30% humidity indicates the presence of dry eye symptoms.[14]

The EASYTEAR view+ is a noninvasive diagnostic tool designed for evaluating the tear film. It plays an essential role in identifying various dry eye syndromes. This device operates as an interferometer, offering detailed insights into the lipid layer of the tear film. It allows practitioners to measure noninvasive tear break-up time after inserting a specialized grid, providing valuable information about tear stability.[15] However, it shares similarities with the Tearscope. The EASYTEAR view+ offers enhanced features, including an integrated stopwatch for precise timing and a light-emitting diode that serves as an effective light source for optimal visualization. This device can operate either as a portable handheld unit or be connected directly to a slit lamp, making it suitable for different clinical environments and practitioner preferences.[15]

Several studies have explored the evaluation of lipid layer patterns (LLP) in smokers, individuals with refractive errors (RE), and those with a high body mass index (BMI), among others.[8,16–20] For example, smokers and individuals with RE or a high BMI tend to exhibit lower LLP values than healthy individuals.[8] However, these investigations often share a common limitation in participant selection: many include only young men or young women, thereby limiting broader population representation. The relationship between TER and LLP has not been systematically studied in smokers, individuals with RE, or those with a high BMI. This research gap highlights the need for further studies to better understand these correlations and their impact on dry eye syndrome.

The integrity of the tear film plays a crucial role in maintaining overall ocular health. Various factors can disrupt this delicate balance, leading to discomfort and dryness. Notably, smoking has been identified as a significant risk factor, as it can adversely affect tear production and quality. Additionally, individuals with RE may experience increased dryness due to imbalance within the visual system. Furthermore, a high BMI has also been linked to dry eye syndrome, as obesity may contribute to inflammation and hormonal changes that affect tear production. Together, these factors highlight the importance of lifestyle modification and appropriate eye care in maintaining a healthy tear film and preventing dry eye symptoms. Therefore, the present comparative observational study examined the correlation between LLP, assessed using EASYTEAR view+, and TER, evaluated using the Delfin VapoMeter, in 3 distinct groups: smokers, individuals with RE, and those with a high BMI.

Objectives

The study aimed to assess the correlation between TER and LLP in smokers, individuals with RE, and those with a high BMI. We hypothesized that smokers and individuals with RE and a high BMI would exhibit thinner lipid layers and higher TER values.

Materials and methods

Subjects

The current comparative observational study included a total of 120 participants aged between 18 and 30 years. They were categorized into 4 distinct groups:

individuals with no ocular diseases or disorders, smokers with a smoking history of 1–11 years, individuals with RE ranging from +7.5 to –7.5 D, and those with a high BMI between 30.0 and 48.7 kg/m^2, with each group comprising 15 women and 15 men. The sample size was determined based on the prevalence of these participant groups among university students.

Before commencing the study, written informed consent was obtained from all participants, ensuring that they were fully informed about the nature of the study. The study received ethical approval from the King Saud University Ethics Committee (Riyadh, Saudi Arabia; approval No. E-24-9087), and all procedures were conducted in accordance with the principles of the Declaration of Helsinki, emphasizing participant safety and rights. The study was conducted at the clinics of the Department of Optometry, College of Applied Medical Sciences, King Saud University (Riyadh, Saudi Arabia).

Specific exclusion criteria were established to maintain the integrity of the study. Individuals at risk for ocular dryness, such as those who had undergone recent eye surgery or were taking certain medications, were excluded. Other risk factors included thyroid disorders, elevated cholesterol levels, diabetes, and deficiencies in vitamins A and D. Furthermore, the study did not include pregnant or breastfeeding women, individuals with eyelid or eyelash abnormalities, or contact lens wearers. Additionally, participants who had undergone thyroid surgery or radioactive iodine therapy within the previous 3 years were excluded. Lastly, individuals with hypertension, anemia, or other serious health conditions were also excluded to ensure a homogeneous study population. The examiner was masked to ensure that she was unaware of the group to which each participant belonged.

The ocular surface disease index

The Ocular Surface Disease Index (OSDI) was used as a valid and reliable tool to assess the severity of dry eye disease and its impact on vision-related function. It is evaluated on a scale from 0 to 100, with higher scores indicating greater severity of ocular dryness. This index effectively differentiates between healthy individuals and patients with dry eye disease, demonstrating both sensitivity and specificity. All participants completed the OSDI at the beginning of the study. A score below 13 was considered indicative of normal ocular health.[21]

EASYTEAR VIEW+

The lipids in the tear film were assessed using the EASYTEAR view+ device (EASYTEAR S.R.L., Trento, Italy). This practical and compact device visually represents interference within the lipid phase of the tear film, providing a clearer understanding of lipid layer thickness (LLT). The same examiner performed all measurements to ensure consistency and reliability. The tests were conducted at a temperature of 20°C and a humidity level below 40%.[8]

Evaluation of the LLP allows detailed assessment of LLT. The LLP was categorized into 5 distinct grades: A, B, C, D, and E, each representing different characteristics and thicknesses of the lipid layer. Grade A showed a subtle gray appearance resembling a very thin white-blue layer associated with a LLT of approx. 13–15 nm. Grade B presented a more compact lipid layer, appearing as a slightly thicker white-blue layer, with an LLT ranging from 30 to 50 nm. Grade C was characterized by dynamic gray waves with a white-blue layer displaying distinct color fluctuations. The thickness of this lipid layer ranged from 50 to 80 nm. Grade D, with an LLT of approx. 80 nm, revealed a dense white-blue layer with a uniform appearance. In contrast, grade E, with an LLT ranging from 90 to 140 nm, displayed a wide variety of colors.[8]

Delfin VapoMeter

The TER was measured using a VapoMeter (Delfin Technologies UK Ltd, Westhumble, UK). The testing procedure was performed 3 times, with 2 readings obtained during each test. The 1st reading was recorded with both eyes open and blinking normally, while the 2nd reading was taken with both eyes closed after a 2-min interval. The TER was calculated by subtracting the reading obtained with the eyes closed from that recorded with the eyes open, and the mean values from the 3 tests were subsequently calculated. Under normal ocular conditions, the TER was considered to be below 25 g/m^2/h, whereas readings above 25 g/m^2/h indicated evaporative dry eye. Additionally, the effect of 2-propanol (70%) pads on the VapoMeter was evaluated.[12,13]

Statistical analyses

The data for this study were collected using Microsoft Excel 2016 (Microsoft Corp., Redmond, USA). Following data collection, a detailed statistical analysis and interpretation of the data were performed using IBM SPSS v. 22.0 (IBM Corp., Armonk, USA) and R v. 4.4.3 (R Foundation for Statistical Computing, Vienna, Austria; session information and package details are listed in Supplementary Table 1).

Normality of age and continuous variables (OSDI and TER) was assessed within each group using the Shapiro–Wilk test (Supplementary Table 2). Due to violations of normality in the healthy, high BMI, and smoker groups, nonparametric methods were employed throughout. Between-group comparisons were performed using the Kruskal–Wallis one-way analysis of variance (ANOVA) by ranks for OSDDI, TER, and LLP. When the omnibus test was significant ($p < 0.05$), post hoc pairwise comparisons were conducted using Dunn's test with Bonferroni adjustment for multiple comparisons (6 pairwise comparisons per

variable). Within-group associations between LLP and continuous variables (OSDI and TER) were assessed using Kendall's tau-b rank correlation coefficient. No multiplicity adjustment was applied to the correlation analyses, as each represented a distinct hypothesis within independent populations. Descriptive statistics are presented as median and interquartile range (IQR) values for age, OSDI, TER, and LLP. Statistical significance was set at $p < 0.05$ for all analyses.

Results

Table 1 presents the median (IQR) values for age, OSDI, TER, and LLP scores in the 4 groups. The OSDI scores varied according to the participants' health conditions. Individuals with healthy eyes had OSDI scores ranging from 1.0 to 13.0, indicating no symptoms of dry eye. In contrast, the study groups showed significantly higher scores and, therefore, a greater degree of discomfort. Subjects with RE had OSDI scores ranging from 2.1 to 50. Participants with a high BMI presented OSDI scores between 8.3 and 72.9. Additionally, smokers had scores ranging from 8.3 to 52.1.

The TER scores also varied across the different groups. In subjects with healthy eyes, TER scores ranged from 5 to 22 g/m^2/h, indicating normal tear evaporation. In individuals with RE, these scores ranged from 4 to 32 g/m^2/h. Participants with a high BMI reported TER scores ranging from 6 to 65 g/m^2/h, while smokers exhibited the widest range, with scores from 5 to 75 g/m^2/h. For LLP, healthy individuals reported grades ranging from 2 to 5. Meanwhile, smokers and individuals with RE and a high BMI reported similar scores ranging from 1 to 5. These findings demonstrate differences in ocular surface health among the groups, highlighting the impact of underlying health conditions on tear film quality parameters.

As shown in Table 2, the Kruskal–Wallis omnibus test for between-group differences was statistically significant for OSDI ($\chi^2 = 62.91$, n = 120, degrees of freedom (df) = 3, $p < 0.001$) and LLP ($\chi^2 = 26.59$, n = 120, df = 3, $p < 0.001$), but not for TER ($\chi^2 = 7.20$, n = 120, df = 3, $p > 0.999$). The corresponding effect size estimates for OSDI and LLP were both large ($\varepsilon^2 = 0.503$ and 0.198, respectively). Results of Dunn's post hoc pairwise comparisons with Bonferroni correction applied for 6 comparisons per variable are shown in Table 3 for OSDI and LLP. For both OSDI and LLP, significant differences ($p < 0.01$ or lower) were observed between the healthy group and each of the 3 study groups. However, there was no evidence of significant differences in these 2 measures among the RE, high BMI, and smoker groups. The analytical results are also illustrated in Fig. 1, which includes side-by-side box-and-whisker plots with individual data points for OSDI, TER, and LLP scores by group.

Results of the within-group correlation analyses are summarized in Fig. 2. Kendall's tau-b (τb) rank-order correlation coefficient was used to quantify the association between LLP scores and both OSDI and TER separately in each of the 4 study groups. No significant associations were detected for OSDI in any of the groups. By contrast, significant negative associations were detected in all groups except the healthy subgroup. For example, in individuals with a high BMI, a strong negative correlation was found between TER and LLP scores ($\tau b = -0.563$, n = 30, $p < 0.001$), indicating that as LLP increased, TER decreased. Similarly, among smokers, a strong negative correlation ($\tau b = -0.457$, $p = 0.002$) was also statistically significant, further underscoring the inverse relationship between these parameters. Subjects with RE exhibited a moderate negative correlation ($\tau b = -0.287$, n = 30, $p = 0.043$) between TER and LLP scores, suggesting a less pronounced but still meaningful inverse relationship. Although the association between TER and LLP scores was also negative among healthy subjects, it was not statistically significant ($\tau b = -0.199$, n = 30, $p = 0.043$).

Table 1. Demographic and clinical characteristics by study group (n = 120; n = 30 per group). Values are presented as median (IQR) for age, OSDI, TER, and LLP scores

Score/group	Healthy (n = 30)	RE (n = 30)	High BMI (n = 30)	Smokers (n = 30)
Age [years]	22.0 (2.8)	21.0 (3.5)	21.0 (2.8)	20.5 (3.0)
OSDI	5.0 (5.5)	22.9 (17.1)	33.3 (21.7)	27.1 (16.1)
TER (g/m^2/h)	12.4 (7.7)	13.6 (12.5)	16.9 (12.6)	15.0 (11.5)
LLP	4.0 (1.0)	2.0 (1.0)	2.5 (3.0)	3.0 (2.0)

Interquartile ranges (IQRs) were calculated using the R function (IQR) and the type 7 method. OSDI – ocular surface disease index; TER – tear evaporation rate; LLP – lipid layer patterns; RE – refractive errors; BMI – body mass index.

Table 2. Kruskal–Wallis omnibus tests for group differences

Variable	χ^2	df	p-value	Sig.
OSDI	62.907	3	0.000	***
TER (g/m^2/h)	7.201	3	0.066	NS
LLP	26.594	3	0.000	***

NS – not significant; df – degrees of freedom; OSDI – ocular surface disease index; TER – tear evaporation rate; LLP – lipid layer patterns; *p < 0.05; **p < 0.01; ***p < 0.001.

Table 3. Dunn's post hoc tests with Bonferroni correction

Variable	Pairwise comparison	Z-statistic	p-value (unadj)	p-value (adj)	Sig.
OSDI	HBMI – Healthy	7.289	0.000	0.000	***
	HBMI – RE	1.935	0.026	0.159	NS
	Healthy – RE	–5.354	0.000	0.000	***
	HBMI – Smoker	1.094	0.137	0.822	NS
	Healthy – Smoker	–6.195	0.000	0.000	***
	RE – Smoker	–0.841	0.200	1.000	NS
TER	Not performed	–	–	–	omnibus NS
LLP	HBMI – Healthy	–3.922	0.000	0.000	***
	HBMI – RE	0.938	0.174	1.000	NS
	Healthy – RE	4.860	0.000	0.000	***
	HBMI – Smoker	–0.938	0.174	1.000	NS
	Healthy – Smoker	2.983	0.001	0.009	**
	RE – Smoker	–1.877	0.030	0.182	NS

Bonferroni correction applied for 6 comparisons per variable. *p < 0.05; ** p < 0.01; ***p < 0.001; NS – not significant. Bonferroni correction was applied for 6 comparisons per variable. *p < 0.05; **p < 0.01; ***p < 0.001; OSDI – ocular surface disease index; TER – tear evaporation rate; LLP – lipid layer patterns; HBMI – high body mass index; RE – refractive errors; unadj – unadjusted; adj – adjusted.

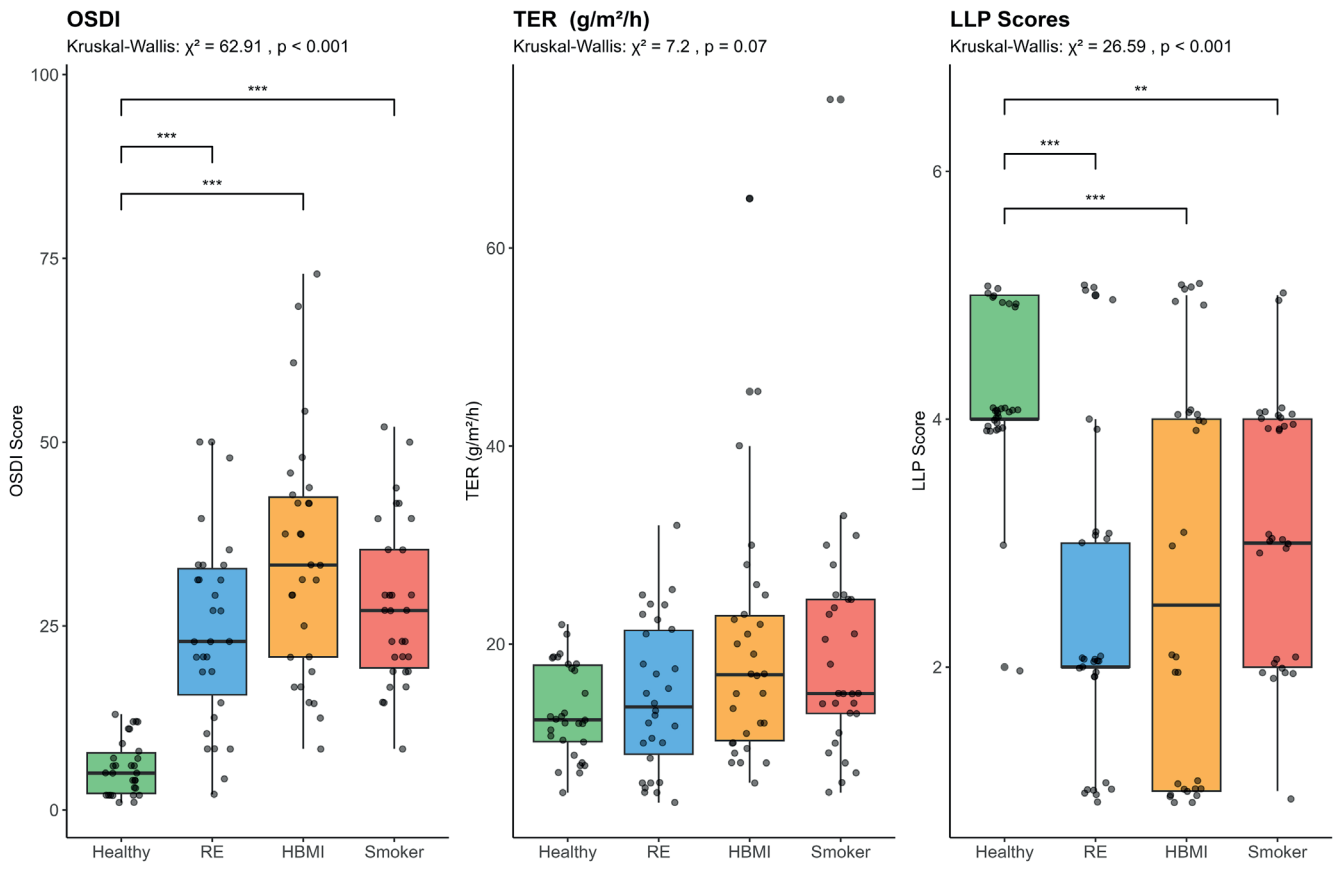


Fig. 1. Box-and-whisker plots with individual data points for ocular surface disease index (OSDI), tear evaporation rate (TER), and lipid layer patterns (LLP) scores by group. Observed values of the Kruskal–Wallis test statistic and unadjusted p-values for the omnibus test of group differences are listed above each panel. Post hoc pairwise comparisons were performed for OSDI and LLP only. Brackets indicate significant differences (see Table 2 for details)

Discussion

It was anticipated that the thickness of the lipid layer would influence TER. An increased LLT was expected to correspond with a reduction in TER.[22] The findings of the current study indeed revealed a robust negative correlation between TER scores and LLP grades, particularly in individuals with RE, a high BMI, and smokers. Previous research has shown that both smokers and individuals with a high BMI tend to have significantly reduced LLT.

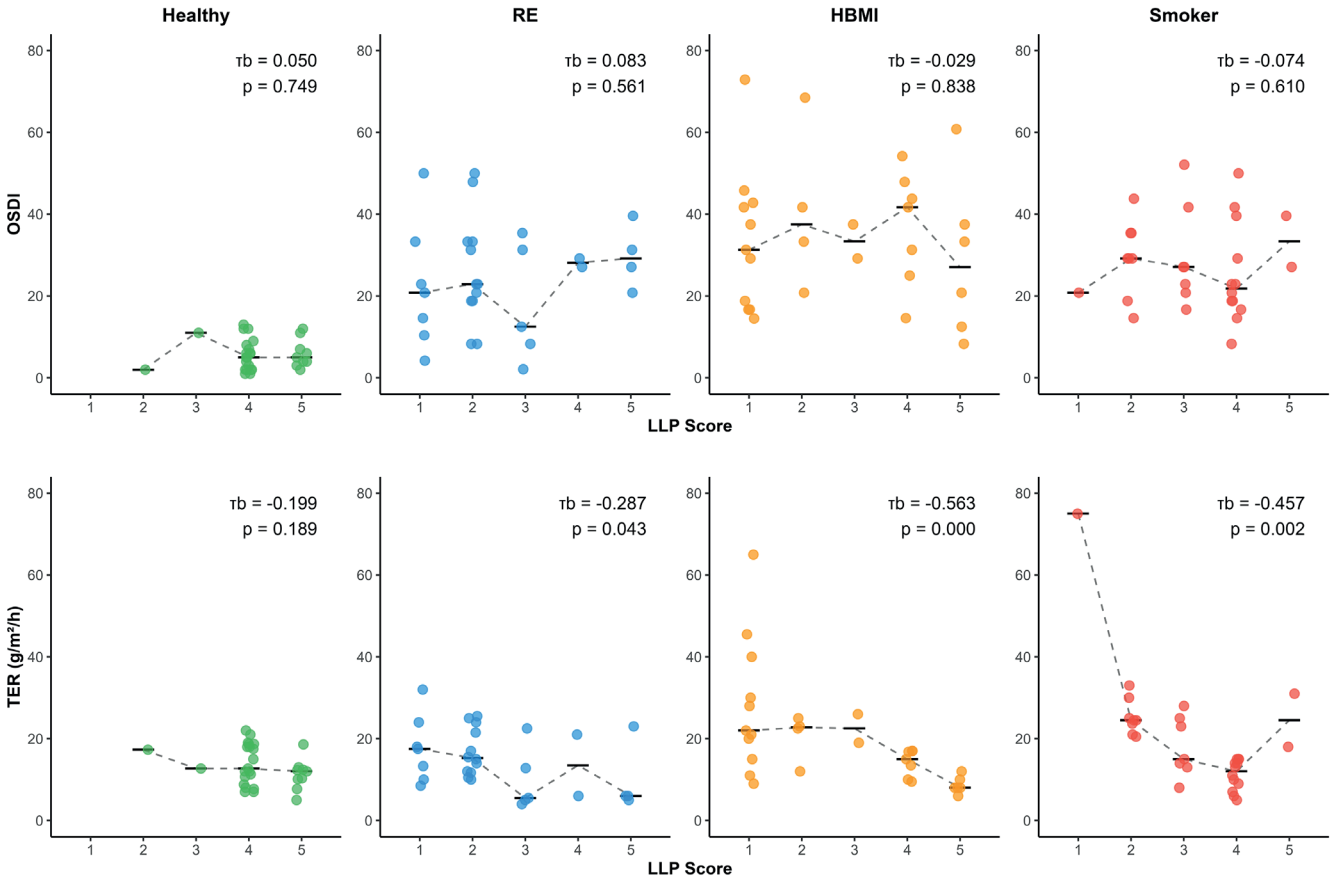


Fig. 2. Relationship between ocular surface disease index (OSDI) and lipid layer patterns (LLP) scores (top row) and between tear evaporation rate (TER) and LLP scores (bottom row) by study group (healthy, refractive errors (RE), high body mass index (BMI), and smokers, from left to right). In each panel, the symbols (filled circles) represent individual data points for n = 30 subjects in each group and have been jittered to minimize overplotting. A short black horizontal line represents the median OSDI or TER score at each LLP level and is connected across LLP values by a dashed line. Annotations in the upper right corner of each panel show the observed value of Kendall's tau-b rank-order correlation coefficient together with the unadjusted p-value. No multiplicity adjustment was applied to the correlation analyses, as each represented a distinct hypothesis within independent populations

Moreover, these groups also demonstrated lower tear meniscus height (TMH) scores compared with the control group.[8]

Elevated TER may increase tear osmolarity, indicating a higher concentration of solutes in the tear film. This phenomenon highlights the importance of understanding the complex relationship between TER and osmolarity, which requires further investigation. In evaporative dry eye, moisture loss through evaporation appears to result primarily from osmotic flow across the corneal and conjunctival surfaces rather than from secretion by the lacrimal glands.[23,24] This process acts as a compensatory mechanism to counteract evaporative moisture loss. As a result, this interaction helps mitigate the expected rise in osmolarity, highlighting the delicate balance within the ocular surface environment.[23,24] Tear evaporation rate plays an essential role in this process and may vary significantly depending on the LLP of the tear film. Importantly, eyes with thinner lipid layers exhibit substantially higher TERs, resulting in increased discomfort and dryness.

The integrity of the tear film lipid layer is essential for maintaining ocular health, and its disruption is particularly significant in smokers, in whom studies indicate an increase in TER of approx. 90–95%.[25,26] The lipid layer acts as a protective barrier, and its thickness is inversely correlated with TER, meaning that as the lipid layer thins, the TER increases. Interestingly, smokers, individuals with conditions such as RE, and those with a high BMI tend to exhibit a reduced thickness of this vital lipid layer. Thinning of the lipid layer disrupts its normal function, which may lead to changes in blinking rate.[27] Blinking plays a vital role in maintaining the tear film by effectively spreading and evenly distributing lipids across its surface. When the blinking rate changes, instability of the tear film occurs, resulting in symptoms associated with dry eye.[28] This instability is primarily driven by increased tear evaporation, causing irritation and discomfort in affected individuals. The interplay between these factors highlights the importance of preserving a healthy tear film for overall ocular health.

Dry eye disease in individuals with meibomian gland dysfunction (MGD) is alarmingly common, with prevalence rates ranging from 20% to 69%.[29] Meibomian gland dysfunction disrupts the delicate balance of the lipid layer, leading to abnormalities in its components that can significantly increase the rate of water evaporation from

the ocular surface. As such, LLT and integrity serve as critical indicators of meibomian gland function. Consequently, evaluation of LLP is an essential diagnostic tool for identifying obstructive MGD. A study involving patients diagnosed with MGD revealed that the thickness of the lipid layer does not consistently correlate with other important tear film parameters, including Schirmer, NIBUT, and TMH scores.[30] This observation has led to the hypothesis that increased tear fluid production may compensate for the deficiency in the lipid layer resulting from MGD.[31] However, the relationship between LLP and tear production remains controversial and warrants further investigation to better understand the underlying mechanisms.

Lipid layer thickness increased with age and exhibited a significant correlation with both meibomian gland secretion and morphology in middle-aged and older individuals with obstructive MGD. Measurement of LLT may serve as a valuable tool for diagnosing MGD. It is crucial to consider age as an important factor when evaluating the implications of LLT measurements.[30] The LLT of patients with dry eye disease should not be considered in isolation; instead, it should be assessed alongside other tear film parameters.[32]

Limitations of the study

The study has several limitations worth mentioning. It involved a relatively small sample size, which may affect the generalizability of the findings. The research focused solely on a specific demographic, namely young individuals, primarily students from King Saud University. Additionally, all participants were recruited from a single geographic region, Riyadh, which may further limit the diversity and scope of the collected data.

Conclusions

Significant differences were found in OSDI and LLP scores between the control group and the 3 study groups (RE, high BMI, and smokers). For TER, a significant difference was observed between the control group and both the RE and smoker groups. Additionally, a significant difference in OSDI scores was identified between the RE and high BMI groups. A strong negative correlation was found between TER and LLP in smokers and individuals with a high BMI. In individuals with RE, a moderate negative correlation was observed between TER and LLP scores.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17036238. The package contains the following files:

Supplementary Table 1. R session information.

Supplementary Table 2. Shapiro–Wilk test for normality by group.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo at https://doi.org/10.5281/zenodo.17036246.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Meznah S. Almutairi https://orcid.org/0000-0003-1802-2403
Basal H. Altoaimi https://orcid.org/0000-0002-2101-7675
Martin Rickert https://orcid.org/0000-0002-7010-7300
Gamal A. El-Hiti https://orcid.org/0000-0001-6675-3126

References

1. Harrell CR, Feulner L, Djonov V, Pavlovic D, Volarevic V. The molecular mechanisms responsible for tear hyperosmolarity-induced pathological changes in the eyes of dry eye disease patients. *Cells*. 2023;12(23):2755. doi:10.3390/cells12232755
2. Stapleton F, Velez FG, Lau C, Wolffsohn JS. Dry eye disease in the young: A narrative review. *Ocul Surf*. 2024;31:11–20. doi:10.1016/j.jtos.2023.12.001
3. Mondal H, Kim HJ, Mohanto N, Jee JP. A review on dry eye disease treatment: Recent progress, diagnostics, and future perspectives. *Pharmaceutics*. 2023;15(3):990. doi:10.3390/pharmaceutics15030990
4. Chan C, Ziai S, Myageri V, Burns JG, Prokopich CL. Economic burden and loss of quality of life from dry eye disease in Canada. *BMJ Open Ophth*. 2021;6(1):e000709. doi:10.1136/bmjophth-2021-000709
5. Sheppard JD, Nichols KK. Dry eye disease associated with meibomian gland dysfunction: Focus on tear film characteristics and the therapeutic landscape. *Ophthalmol Ther*. 2023;12(3):1397–1418. doi:10.1007/s40123-023-00669-1
6. Doctor MB, Basu S. Lacrimal gland insufficiency in aqueous deficiency dry eye disease: Recent advances in pathogenesis, diagnosis, and treatment. *Semin Ophthalmol*. 2022;37(7–8):801–812. doi:10.1080/08820538.2022.2075706
7. Sakakura S, Inagaki E, Ochiai Y, et al. A comprehensive assessment of tear-film-oriented diagnosis (TFOD) in a dacryoadenectomy dry eye model. *Int J Mol Sci*. 2023;24(22):16510. doi:10.3390/ijms242216510
8. Fagehi R, El-Hiti GA, Almojalli A, et al. Assessment of tear film parameters in smokers and subjects with a high body mass index. *Optom Vis Sci*. 2022;99(4):358–362. doi:10.1097/OPX.0000000000001891
9. Aldarwesh A, Fagehi R, Alruways K, et al. Assessment of tear film parameters post-treatment with commercial eyelid cleaning wipes: A pilot study. *Int J Ophthalmol*. 2024;17(4):659–664. doi:10.18240/ijo.2024.04.08
10. Papas EB. Diagnosing dry-eye: Which tests are most accurate? *Contact Lens Anterior Eye*. 2023;46(5):102048. doi:10.1016/j.clae.2023.102048
11. Kulovesi P, Rantamäki AH, Holopainen JM. Surface properties of artificial tear film lipid layers: Effects of wax esters. *Invest Ophthalmol Vis Sci*. 2014;55(7):4448. doi:10.1167/iovs.14-14122
12. Abusharha AA, Al Yami A, Alsreea K, et al. Repeatability and reproducibility of tear film evaporation rate measurement using a new closed-chamber evaporimeter. *Open Ophthalmol J*. 2021;15(1):117–121. doi:10.2174/1874364102115010117
13. Kim YH, Graham AD, Li W, et al. Tear-film evaporation flux and its relationship to tear properties in symptomatic and asymptomatic soft-contact-lens wearers. *Contact Lens Anterior Eye*. 2023;46(4):101850. doi:10.1016/j.clae.2023.101850
14. Tsubota K, Yamada M. Tear evaporation from the ocular surface. *Invest Ophthalmol Vis Sci*. 1992;33(10):2942–2950. PMID:1526744.

15. Bandlitz S, Peter B, Pflugi T, et al. Agreement and repeatability of four different devices to measure non-invasive tear breakup time (NIBUT). *Contact Lens Anterior Eye*. 2020;43(5):507–511. doi:10.1016/j.clae.2020.02.018
16. Kim JS, Lee H, Choi S, Kim EK, Seo KY, Kim TI. Assessment of the tear film lipid layer thickness after cataract surgery. *Semin Ophthalmol*. 2018;33(2):231–236. doi:10.1080/08820538.2016.1208764
17. Maïssa C, Guillon M. Tear film dynamics and lipid layer characteristics: Effect of age and gender. *Contact Lens Anterior Eye*. 2010;33(4):176–182. doi:10.1016/j.clae.2010.02.003
18. Sabucedo-Villamarin B, Garcia-Queiruga J, Pena-Verdeal H, Garcia-Resua C, Yebra-Pimentel E, Giraldez MJ. Diagnostic cut-off values based on lipid layer pattern for dry eye disease subtypes assessment. *J Clin Med*. 2025;14(2):623. doi:10.3390/jcm14020623
19. Yazdani M. Tear film lipid layer and corneal oxygenation: A new function? *Eye (Lond)*. 2023;37(17):3534–3541. doi:10.1038/s41433-023-02557-1
20. Weng HY, Ho WT, Chiu CY, Tsai TY, Chang SW. Characteristics of tear film lipid layer in young dry eye patients. *J Formos Med Assoc*. 2021;120(7):1478–1484. doi:10.1016/j.jfma.2020.10.028
21. Schiffman RM. Reliability and validity of the Ocular Surface Disease Index. *Arch Ophthalmol*. 2000;118(5):615. doi:10.1001/archopht.118.5.615
22. King-Smith PE, Hinel EA, Nichols JJ. Application of a novel interferometric method to investigate the relation between lipid layer thickness and tear film thinning. *Invest Ophthalmol Vis Sci*. 2010;51(5):2418. doi:10.1167/iovs.09-4387
23. Levin MH, Verkman AS. Aquaporin-dependent water permeation at the mouse ocular surface: In vivo microfluorimetric measurements in cornea and conjunctiva. *Invest Ophthalmol Vis Sci*. 2004;45(12):4423. doi:10.1167/iovs.04-0816
24. Craig JP, Tomlinson A. Importance of the lipid layer in human tear film stability and evaporation. *Optom Vis Sci*. 1997;74(1):8–13. doi:10.1097/00006324-199701000-00014
25. Abusharha A, Alturki AA, Alanazi SA, et al. Assessment of tear-evaporation rate in thyroid-gland patients. *Clin Ophthalmol*. 2019;13:131–135. doi:10.2147/OPTH.S188614
26. Alanazi AS, Abusharha AA, Fagehi R, et al. Assessment of the tear evaporation rate in chronic smokers using Delfin VapoMeter. *Int J Ophthalmol Vis Sci*. 2019;4(2):37. doi:10.11648/j.ijovs.20190402.12
27. Li Y, Li S, Zhou J, Liu C, Xu M. Relationship between lipid layer thickness, incomplete blinking rate and tear film instability in patients with different myopia degrees after small-incision lenticule extraction. *PLoS One*. 2020;15(3):e0230119. doi:10.1371/journal.pone.0230119
28. Wang MTM, Tien L, Han A, et al. Impact of blinking on ocular surface and tear film parameters. *Ocul Surf*. 2018;16(4):424–429. doi:10.1016/j.jtos.2018.06.001
29. Schaumberg DA, Nichols JJ, Papas EB, Tong L, Uchino M, Nichols KK. The International Workshop on Meibomian Gland Dysfunction: Report of the Subcommittee on the Epidemiology of, and Associated Risk Factors for MGD. *Invest Ophthalmol Vis Sci*. 2011;52(4):1994. doi:10.1167/iovs.10-6997e
30. Li J, Ma J, Hu M, Yu J, Zhao Y. Assessment of tear film lipid layer thickness in patients with meibomian gland dysfunction at different ages. *BMC Ophthalmol*. 2020;20(1):394. doi:10.1186/s12886-020-01667-8
31. Arita R, Morishige N, Koh S, et al. Increased tear fluid production as a compensatory response to meibomian gland loss. *Ophthalmology*. 2015;122(5):925–933. doi:10.1016/j.ophtha.2014.12.018
32. Lee Y, Hyon JY, Jeon HS. Characteristics of dry eye patients with thick tear film lipid layers evaluated by a LipiView II interferometer. *Graefes Arch Clin Exp Ophthalmol*. 2021;259(5):1235–1241. doi:10.1007/s00417-020-05044-5

Prediction model for postoperative urinary tract infection after unilateral pyeloplasty in children

Hongyang Wang[1,B–F], Chunsheng Hao[1,B,C,E,F], Long Li[2,B,E,F], Qing Sun[1,B,E,F], Xiaomeng Cui[1,E,F], Dongsheng Bai[1,A,E,F], Jinqiu Song[1,A–F]

[1] Capital Center for Children's Health, Capital Medical University, Beijing, China
[2] Department of Pediatric Surgery, Capital Institute of Pediatric, Research Unit of Minimally Invasive Pediatric Surgery on Diagnosis and Treatment, Chinese Academy of Medical Sciences, Beijing, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1017–1025

Address for correspondence
Jinqiu Song
E-mail: songjingiu@163.com

Funding sources
This study was supported by the Research Foundation of the Capital Institute of Pediatrics (grant No. LCYJ-2023-12), the Research Unit of Minimally Invasive Pediatric Surgery on Diagnosis and Treatment, Chinese Academy of Medical Sciences (grant No. 2021RU015), and the Beijing Municipal Administration of Hospitals Incubating Program (grant No. PX2025046).

Conflict of interest
None declared

Received on February 13, 2025
Reviewed on March 8, 2025
Accepted on August 20, 2025

Published online on June 23, 2026

Abstract

Background. Postoperative urinary tract infection (UTI) following pyeloplasty remains a significant complication and continues to pose challenges in pediatric urological care.

Objectives. This study aimed to develop a simplified predictive model to identify risk factors for postoperative UTI after unilateral pyeloplasty and to support clinicians in implementing preventive strategies targeting modifiable risk factors.

Materials and methods. Clinical data from children who underwent unilateral pyeloplasty at the Children's Hospital of Capital Institute of Pediatrics (Beijing, China) between January 2012 and January 2022 were retrospectively analyzed. Variables including sex, age, body mass index (BMI), surgical modality, drainage tube type, and parameters from blood and urine tests were evaluated. Statistical analyses, including least absolute shrinkage and selection operator (LASSO) regression, logistic regression, and random forest modeling, were performed to identify significant predictive factors. Variables with the greatest predictive importance were used to develop a nomogram, and its clinical utility was evaluated using decision curve analysis (DCA).

Results. Among 764 patients, 265 (35%) developed postoperative UTI. Key risk factors included surgical modality, laterality of ureteropelvic junction obstruction (UPJO), drainage tube type, blood urea nitrogen (BUN) level, and patient height. LASSO regression identified 14 predictive variables, while logistic regression determined independent risk and protective factors. Ultimately, 8 variables (e.g., sex, operative time, drainage tube type, history of infection, history of fistula, age, BUN level, and renal cortical thickness) were selected for development of the nomogram predicting postoperative UTI risk after unilateral pyeloplasty.

Conclusions. This study identified 8 factors associated with postoperative UTI following unilateral pyeloplasty in children. The developed predictive model may assist clinicians in identifying high-risk patients, thereby supporting improved perioperative planning and postoperative management.

Key words: risk factors, urinary tract infection, prediction model, unilateral pyeloplasty

Cite as
Wang H, Hao C, Li L, et al. Prediction model for urinary tract infection after unilateral pyeloplasty in children. *Adv Clin Exp Med*. 2026;35(6):1017–1025. doi:10.17219/acem/209761

DOI
10.17219/acem/209761

Highlights

- Postoperative urinary tract infection (UTI) occurred in 35% of children after unilateral pyeloplasty, highlighting a major complication in pediatric urology.
- Eight key predictors were identified – gender, operative time, drainage tube type, infection history, fistula history, age, blood urea nitrogen (BUN) level, and renal cortex thickness.
- A nomogram-based predictive model was developed, demonstrating strong clinical utility through decision curve analysis (DCA).
- Early risk stratification enables targeted prevention, supporting optimized preoperative planning and postoperative care for high-risk children.

Background

Ureteropelvic junction obstruction (UPJO) is defined as an obstruction at the junction between the kidney and ureter, resulting in decreased urine flow from the renal pelvis to the ureter. Ureteropelvic junction obstruction is one of the main causes of infantile hydronephrosis. If left untreated, it may lead to hydronephrosis, chronic infection, or urolithiasis and may ultimately result in progressive renal insufficiency.[1] Ureteropelvic junction obstruction may be either congenital or acquired, and endogenous and exogenous risk factors include urolithiasis, postoperative, inflammatory, or ischemic stenosis, fibroepithelial polyps, adhesions, and malignant tumors.

A variety of treatment options are available for UPJO; however, surgery remains the gold standard, particularly when urinary drainage does not improve after 18 months of age. Surgical management of UPJO has advanced considerably over recent decades, evolving from open pyeloplasty (OP) to laparoscopic pyeloplasty (LP) and robotic-assisted laparoscopic pyeloplasty (RALP). Although minimally invasive approaches have gained increasing acceptance in the treatment of UPJO, OP as originally described by Anderson and Hynes, remains the standard surgical treatment, with reported long-term success rates exceeding 90%.[2] Compared with traditional OP, LP offers advantages such as faster recovery, shorter hospitalization, and fewer complications, whereas OP is often associated with shorter operative time.[3] However, some studies have reported shorter operative times for LP compared with OP.[4] Indications for surgical treatment of UPJO include worsening hydronephrosis during surveillance, patient-reported symptoms such as flank pain, recurrent urinary tract infection (UTI), persistent or poorly controlled hypertension, and reduced or declining differential renal function, typically defined as ipsilateral renal function <40% on diuretic renography.[5,6]

However, despite its high success rate, surgical treatment of UPJO is not without complications. The main postoperative complications include UTI, urinary extravasation or leakage, pyelonephritis, bleeding, and recurrent UPJO.[7] Among these, postoperative UTI remains a frequent and clinically challenging complication for pediatric urologists.

Urinary tract infection, most commonly caused by *Escherichia coli*, is one of the most common bacterial infections in children.[8] Its incidence is higher in boys than in girls during the 1st year of life (3.7% vs 2.0%), whereas after infancy it becomes more common in girls.[9] Previous studies have identified several risk factors for pediatric UTI, including male sex, body weight, elevated blood urea nitrogen (BUN) levels, recurrent UTI within 3 months, prolonged catheter retention, double-J stent placement, and bilateral double-J stent retention.[10,11] However, relatively few studies have specifically investigated risk factors for UTI following unilateral pyeloplasty, particularly in large pediatric cohorts.

Objectives

We aimed to develop a predictive model for UTI after unilateral pyeloplasty in children at a large Chinese center, which may help reduce the incidence of postoperative UTI and provide a reliable reference for the management of UTI in children after unilateral pyeloplasty.

Materials and methods

Patients

Clinical data of children who underwent unilateral pyeloplasty at the Children's Hospital of the Capital Institute of Pediatrics (Beijing, China) between January 2012 and January 2022 were retrospectively analyzed. Patient data were retrieved in May 2023. Inclusion criteria were as follows: 1) children undergoing unilateral pyeloplasty for the first time; 2) patients meeting the diagnostic criteria for UPJO according to the European Association of Urology (EAU) guidelines; and 3) patients with complete follow-up data. Exclusion criteria were as follows: 1) patients with bilateral hydronephrosis; 2) patients who received a double-J stent (internal drainage) and temporary pyelostomy fistula (external drainage) simultaneously; 3) patients with multiple congenital urinary tract strictures; 4) patients with duplicated

kidneys or double ureters; 5) patients whose parients or legal guardians refused participation in the study or declined to sign the informed consent form; and 6) patients with incomplete follow-up data. The study protocol was approved by the Institutional Review Board of the Children's Hospital of the Capital Institute of Pediatrics (approval No. SHIERILM2025005) before commencement of the study, and all patients' guardians signed informed consent forms. Written informed consent for participation in the study and publication of potentially identifiable images or data was obtained from all parents.

Variables

Variables including sex, age, height, weight, body mass index (BMI), surgical modality, indwelling drainage tube type (double-J stent or pyelostomy fistula), serum creatinine (Cr), BUN, estimated glomerular filtration rate (eGFR), neutrophil percentage (N%), lymphocyte percentage (L%), neutrophil-to-lymphocyte ratio (N/L), anterior-posterior diameter of the renal pelvis (APD), caliectasis, and renal cortical thickness were analyzed in this study.

Postoperative UTI definition

Postoperative UTI was defined according to the 2015 EAU guidelines on UTI[12] and the 2016 evidence-based guidelines on UTI issued by the Urology Branch of the Chinese Medical Association.[13] Specifically, postoperative UTI included patients presenting with symptoms such as fever, turbid urine, abnormal urine color, hematuria, urinary retention, urinary incontinence, urinary frequency, urinary urgency, dysuria, or interrupted urination. In addition, the definition required 2 consecutive routine urine test results showing a white blood cell (WBC) count >5/high-power field (HPF). Furthermore, at least 1 of the following criteria had to be fulfilled: 1) urine culture showing $>10^3$ colony-forming units (CFU)/mL of uropathogens in patients with urinary catheters; or 2) urine culture showing $>10^5$ CFU/mL of uropathogens regardless of catheter use. Asymptomatic bacteriuria was defined as the presence of a positive urine culture in patients without symptoms, accompanied by a urine test showing a WBC count >5/HPF.

Preventative measures

All patients received cefuroxime sodium (30 mg/kg/day divided into 3–4 doses) and ceftriaxone sodium (20–80 mg/kg/day) for 2 days before surgery. In addition, all patients received postoperative intravenous anti-inflammatory treatment. A Foley catheter was retained during surgery and removed 3–5 days postoperatively. A double-J stent was placed during surgery and removed 8 weeks postoperatively. In this study, postoperative UTI was defined as UTI occurring within 8 weeks after surgery.

Statistical analyses

The statistical analysis was carried out using R v. 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria) and IBM SPSS v. 23.0 (IBM Corp., Armonk, USA). Prior to model construction, comprehensive assumption checks were performed. Variance inflation factors (VIFs) were calculated for all variables, with generalized VIF values for categorical variables ranging from 1.12 to 3.47 (Supplementary Table 1), confirming the absence of multicollinearity (VIF < 5). Box–Tidwell tests demonstrated linearity in the logit for continuous variables ($p = 0.213–0.586$), and Shapiro–Wilk tests confirmed normality for all continuous variables ($p > 0.05$). Results of these tests are presented in Supplementary Table 1. Continuous variables with skewed distributions are presented as median [Q1–Q3], whereas normally distributed variables (e.g., albumin) are shown as mean ± standard deviation (SD).

Participants were randomly divided into training (70%) and validation (30%) datasets. The random seed was set to 20230815 (set.seed(20230815) in R) prior to data partitioning to ensure reproducibility. The least absolute shrinkage and selection operator (LASSO) regression was applied exclusively for variable selection. Ten-fold cross-validation was used to optimize the penalty parameter (λ), with the optimal λ ($\lambda\_min = 0.032$) selected according to the minimum partial likelihood deviance. LASSO model settings included a maximum of 1,000 iterations and a convergence tolerance of 1×10^{-4}. All 23 candidate variables (listed in Table 1) were included, and variables with non-zero coefficients were retained.

Variables selected with LASSO were entered into a multivariable logistic regression model. Univariable analysis was intentionally omitted to avoid biased variable selection and contradictory results. Model discrimination was assessed using the area under the receiver operating characteristic (ROC) curve (AUC). Calibration was evaluated using the Hosmer–Lemeshow test and calibration plots with 1,000 bootstrap resamples. Clinical utility was quantified through decision curve analysis (DCA) across probability thresholds ranging from 0% to 100%. Post hoc power analysis using G*Power 3.1 demonstrated 98.7% power to detect an odds ratio (OR) >2.0 at $\alpha = 0.05$. Bootstrap validation using 1,000 resamples showed consistent OR estimates within ±5% of the original values.

Results

Patients' demographic and clinical characteristics

A total of 764 patients were included, of whom 265 (34.7%) developed postoperative UTI. The UTI group consisted predominantly of males (86.0% vs 79.4% in the non-UTI group) and younger children (median age: 7.3 months

Table 1. Comparison of baseline characteristics between the urinary tract infection (UTI) and non-UTI groups by univariate analysis

Variable		Overall (n = 764)	Non-UTI (n = 499)	UTI (n = 265)
Gender, n (%)	female	140 (18.3)	103 (20.6)	37 (14.0)
	male	624 (81.7)	396 (79.4)	228 (86.0)
Surgery modality, n (%)	OP	419 (54.8)	310 (62.1)	109 (41.1)
	LP	345 (45.2)	189 (37.9)	156 (58.9)
Operation time, median (Q1–Q3) [min]		90 (70–120)	90 (70–120)	94 (75–120)
Drainage tube type, n (%)	pyelostomy fistula	334 (43.7)	279 (55.9)	55 (20.8)
	double J-stent	430 (56.3)	220 (44.1)	210 (79.2)
Number of operations, n (%)	1	26 (3.4)	15 (3.0)	11 (4.2)
	2	738 (96.6)	484 (97.0)	254 (95.8)
History of UTI, n (%)	no	722 (94.5)	478 (95.8)	244 (92.1)
	yes	42 (5.5)	21 (4.2)	21 (7.9)
History of pyelostomy fistula, n (%)	no	740 (96.9)	495 (99.2)	245 (92.5)
	yes	24 (3.1)	4 (0.8)	20 (7.5)
Abdominal pain, n (%)	no	621 (81.3)	377 (75.6)	244 (92.1)
	yes	143 (18.7)	122 (24.4)	21 (7.9)
Cause of obstruction, n (%)	congenital	742 (97.1)	484 (97.0)	258 (97.4)
	vascular	9 (1.2)	7 (1.4)	2 (0.8)
	polyp	9 (1.2)	7 (1.4)	2 (0.8)
	valve	2 (0.3)	1 (0.2)	1 (0.4)
	high-positioned ureter	1 (0.1)	0 (0.0)	1 (0.4)
	ureteral rupture	1 (0.1)	0 (0.0)	1 (0.4)
Laterality, n (%)	left	555 (72.6)	374 (74.9)	181 (68.3)
	right	209 (27.4)	125 (25.1)	84 (31.7)
Age, median (Q1–Q3) [months]		23 (6.4–60)	36 (11–72)	7.3 (3.9–24)
Weight, median (Q1–Q3) [kg]		12.5 (8.5–20)	16 (10–23)	9.2 (7.5–13.3)
Height, median (Q1–Q3) [cm]		85 (68–114)	100 (74–120)	71 (65–88)
BMI, median (Q1–Q3) [km/m^2]		17.17 (15.5–19.1)	17.01 (15.3–18.9)	17.52 (16.0–19.3)
TP, median (Q1–Q3) [g/L]		65.9 (61.5–70.7)	67.8 (63.2–71.9)	62.3 (58.8–67.2)
Albumin, median (Q1–Q3) [g/L]		43.6 (41.7–45.6)	43.7 (41.9–45.5)	43.5 (41.4–45.9)
GLB, median (Q1–Q3) [g/L]		22.4 (17.8–27.3)	24.2 (20.0–28.1)	18.2 (16.0–22.7)
A/G ratio, median (Q1–Q3)		1.99 (1.61–2.45)	1.81 (1.56–2.21)	2.40 (1.92–2.74)
UA, median (Q1–Q3) [μmol/L]		268 (224–308)	271 (230–311)	260 (216–305)
Cr, median (Q1–Q3) [μmol/L]		29.0 (22.6–37.9)	31.4 (25.0–39.5)	23.6 (20.4–30.7)
BUN, median (Q1–Q3) [mmol/L]		4.10 (2.80–5.10)	4.30 (3.20–5.40)	3.20 (2.29–4.52)
eGFR, median (Q1–Q3) [mL/min/m^2]		36.86 (8.77–61.30)	49.13 (24.88–67.63)	10.29 (4.90–35.57)
Neutrophil ratio (N%), median (Q1–Q3)		33.0% (22.1–45.3)	37.0% (25.6–46.7)	26.5% (19.0–41.5)
Lymphocyte ratio (L%), median (Q1–Q3)		57.0% (44.9–67.6)	53.0% (42.7–63.8)	62.8% (48.0–71.0)
N/L ratio, median (Q1–Q3)		0.59 (0.33–1.02)	0.71 (0.41–1.10)	0.43 (0.27–0.88)
APD, median (Q1–Q3) [cm]		3.30 (2.70–4.00)	3.30 (2.65–3.80)	3.40 (2.70–4.30)
Renal cortical thickness, median (Q1–Q3) [mm]		2.00 (1.60–3.00)	2.40 (2.00–3.40)	1.90 (1.30–2.50)

BMI – body mass index; TP – total protein; Cr – creatinine; UA – uric acid; BUN – blood urea nitrogen; eGFR – estimated glomerular filtration rate; A/G – albumin/globulin ratio; GLB – globulin; N – neutrophil percentage; L – lymphocyte percentage; N/L – neutrophil-to-lymphocyte ratio; APD – anterior–posterior diameter; RPT – renal parenchymal thickness; OP – open pyeloplasty; LP – laparoscopic pyeloplasty; UTI – urinary tract infection.

[3.9–24.0] vs 36.0 months [11.0–72.0]). Laparoscopic pyeloplasty was performed more frequently in patients with UTI (58.9% vs 37.9%), whereas OP was more common in the non-UTI group (62.1% vs 41.1%). Double-J stent use was markedly higher in the UTI group (79.2% vs 44.1%), and renal cortical thickness was lower (median: 1.9 mm [1.3–2.5] vs 2.4 mm [2.0–3.4]). Key biochemical differences were also observed. Patients with UTI showed lower median BUN levels (3.2 mmol/L [2.3–4.5] vs 4.3 mmol/L [3.2–5.4]) and eGFR values (10.29 mL/min/m^2 [4.9–35.6]

vs 49.13 mL/min/m^2 [24.9–67.6]). The N/L ratio was lower in UTI cases (0.43 [0.27–0.88] vs 0.71 [0.41–1.10]), whereas the L% was higher (62.8% [48.0–71.0] vs 53.0% [42.7–63.8]). No notable differences were observed in albumin levels (43.5 g/L [41.4–45.9] vs 43.7 g/L [41.9–45.5]) or renal pelvis APD (3.4 cm [2.7–4.3] vs 3.3 cm [2.7–3.8]). Full baseline characteristics are presented in Table 1.

Identification of predictive variables for UTI after unilateral UPJO through LASSO regression analysis

In the present study, the LASSO regression method was employed to identify significant predictive variables. By applying a penalty parameter (λ) and using 10-fold cross-validation, an optimal set of variables was selected. At the minimum value of log(λ), denoted as log(λ)_min, a random seed was set to ensure reproducibility of the results. The final model included 14 predictive variables, with coefficients presented in Table 1. The intercept coefficient was –0.610070252. The coefficient for sex (gender1) was 0.115231168, suggesting that changes in this variable were associated with changes in the risk of postoperative UTI following unilateral UPJO, while other variables remained constant. The coefficient for operative time was 0.003582231, indicating a slight increase in postoperative UTI risk with longer operative duration. The coefficient for drainage type (drainage1, 1.077044923) was the highest among all variables, indicating a strong association with postoperative UTI risk. The coefficients for infection history and fistula history were 0.351245843 and 0.71021905, respectively, suggesting positive associations with postoperative UTI risk. LASSO coefficients are regularized estimates for variable selection; their magnitude differs from logistic regression coefficients. Direction and significance were confirmed in the final model.

The coefficients for age and height were –0.00270678 and –0.004534245, respectively, indicating negative associations with postoperative UTI risk, with age showing a relatively small effect size. Other biochemical indicators, including TP, albumin/globulin (A/G) ratio, Cr, BUN, and eGFR, also had negative coefficients, suggesting that higher values were associated with lower postoperative UTI risk. The coefficient for APD was 0.006521787, indicating a positive association with postoperative UTI risk, whereas the coefficient for renal cortical thickness was –0.152582245, suggesting that thinner renal cortex was associated with increased postoperative UTI risk. Further results are presented in Table 2 and Fig. 1.

Variable selection and multivariable logistic regression analysis

Employing LASSO regression with 10-fold cross-validation, 8 predictive variables were selected from 23 candidate features (Fig. 1A,B). These variables were subsequently included in a multivariable logistic regression model. Double-J stent placement emerged as the strongest risk factor (OR = 6.41, 95% confidence interval (95% CI): 3.78–10.89, p < 0.001), followed by male sex (OR = 2.15, 95% CI: 1.28–3.63, p = 0.004) and history of fistula (OR = 4.26, 95% CI: 1.85–9.81, p = 0.001). Protective factors included age (OR = 0.98 per month, 95% CI: 0.97–0.99, p < 0.001), BUN level (OR = 0.85 per mmol/L, 95% CI: 0.74–0.98, p = 0.030), and renal cortical thickness (OR = 0.77 per mm, 95% CI: 0.61–0.96, p = 0.023). Variables such as A/G ratio and APD did not reach statistical significance (p > 0.05). Complete results are presented in Table 2.

Development, performance, and internal validation of the predictive nomogram

Utilizing the 8 identified predictive factors, a nomogram was developed to assess the risk of UTI after unilateral UPJO, as shown in Fig. 2. The nomogram assigned scores to each variable, and the total score was calculated as the sum of the individual scores for all 8 variables. At the bottom of Fig. 2, a predictive scale correlates total scores with the corresponding probabilities of postoperative UTI after unilateral UPJO. An increase in the total score indicates a higher risk of postoperative UTI.

Table 2. Variables selected with LASSO regression and final multivariable logistic regression results

Variable	LASSO coefficient	Adjusted OR (95% CI)	p-value
Intercept	–0.6101	–	–
Double-J stent (yes vs no)	1.0770	6.41 (3.78–10.89)	<0.001
Male sex	0.1152	2.15 (1.28–3.63)	0.004
Operative time (per min)	0.0036	1.01 (1.00–1.02)	0.003
History of UTI	0.3512	3.93 (1.24–12.45)	0.020
History of pyelostomy fistula	0.7102	4.26 (1.85–9.81)	0.001
Age (per month)	–0.0027	0.98 (0.97–0.99)	<0.001
BUN (per mmol/L)	–0.0106	0.85 (0.74–0.98)	0.030
Renal cortical thickness (per mm)	–0.1526	0.77 (0.61–0.96)	0.023

LASSO – least absolute shrinkage and selection operator; OR – odds ratio; 95% CI – 95% confidence interval; BUN – blood urea nitrogen; UTI – urinary tract infection.

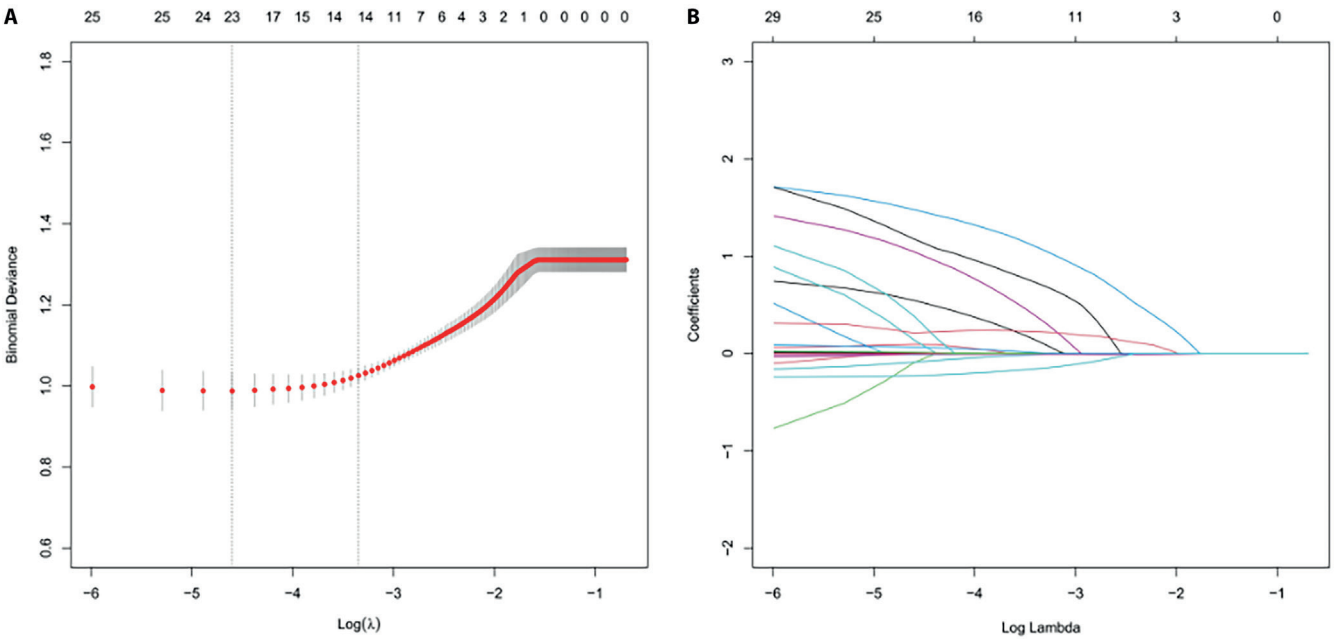


Fig. 1. Least absolute shrinkage and selection operator (LASSO) regression process for variable selection. A. Optimal parameter (λ) determination using 10-fold cross-validation, where the dotted vertical lines indicate the optimal points based on the minimum criteria and the one-standard-deviation rule; B. LASSO coefficient profiles for candidate variables

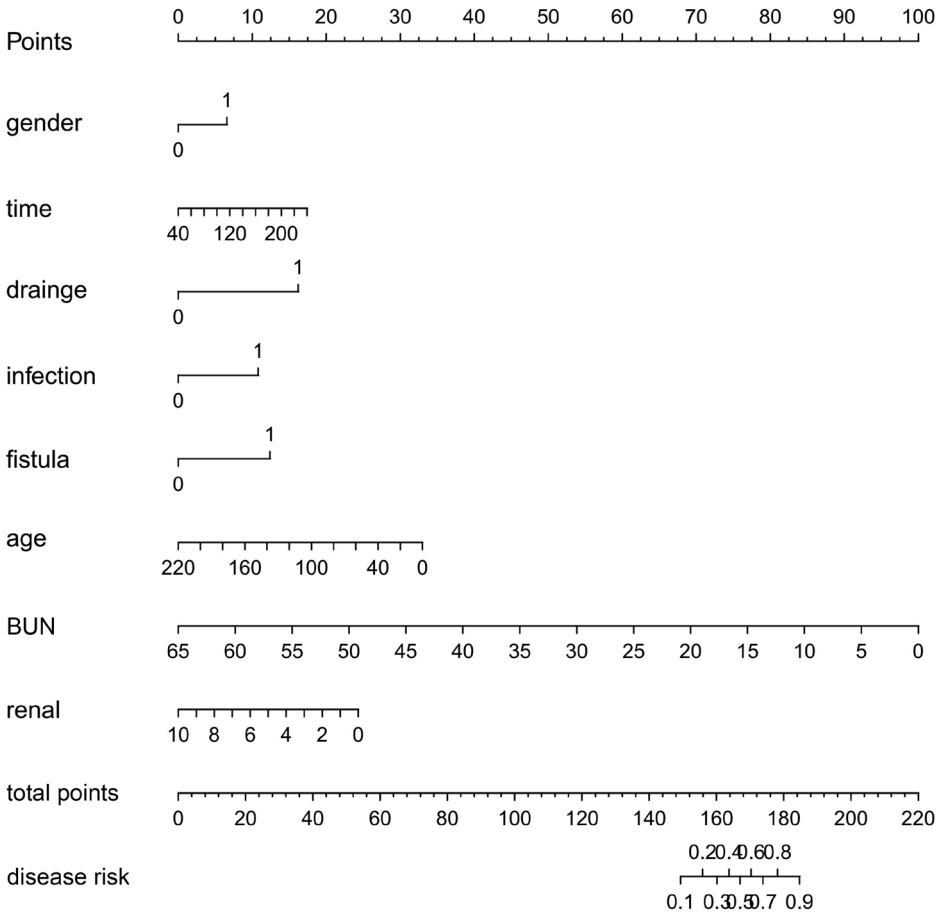


Fig. 2. Proposed nomogram for predicting urinary tract infection (UTI) after unilateral ureteropelvic junction obstruction (UPJO) surgery

The nomogram demonstrated AUC values of 0.844 (95% CI: 0.808–0.879) and 0.789 (95% CI: 0.723–0.855) in the training and validation datasets, respectively (Fig. 3A). Calibration plots for both cohorts closely aligned with the ideal curve after bias correction, indicating good agreement between predicted probabilities and observed outcomes (Fig. 3B,C). Furthermore, the Hosmer–Lemeshow test yielded a p-value of 0.666, supporting adequate

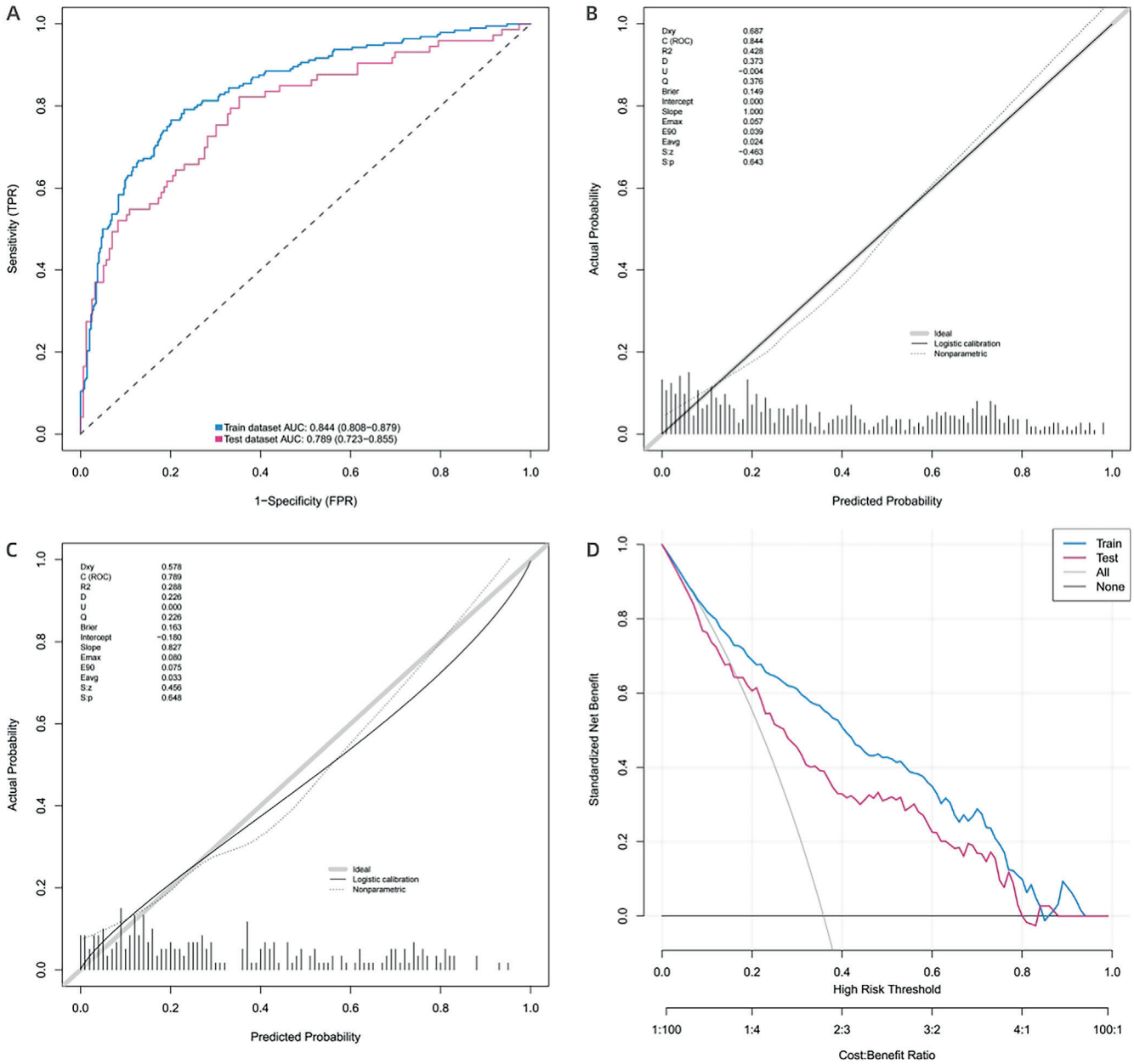


Fig. 3. A. Receiver operating characteristic (ROC) curves for the proposed nomogram; B,C. Calibration curves of the nomogram for the training dataset (B) and internal validation dataset (C); D. Decision curve analysis (DCA) of the proposed nomogram

model calibration. The DCA demonstrated significant net clinical benefit across threshold probabilities ranging from 9% to 89% (Fig. 3D). Collectively, these findings support the discriminative performance, calibration, clinical utility, and potential generalizability of the nomogram.

Discussion

Ureteropelvic junction obstruction is defined as a blockage at the junction between the renal pelvis and the ureter, with an estimated incidence of 1 per 1,000–1,500 individuals.[7] It is more common in boys than in girls and represents the most frequent cause of antenatally detected hydronephrosis, accounting for approx. 80% of cases.[14] At present, surgical treatment remains the gold standard for UPJO, with reported long-term success rates exceeding 90%. Surgical management of UPJO has advanced considerably over recent decades, from the introduction of OPin 1892[15] to the development of LP in 1993,[16] followed by RALP in 2002.[17]

Despite the high success rate of pyeloplasty, postoperative complications remain clinically important. The 2 most common complications include UTI and prolonged urinary leakage. Among these, postoperative UTI remains a frequent and challenging complication in pediatric urology. Therefore, identifying children at high risk of postoperative UTI before pyeloplasty is important for implementing preventive strategies targeting modifiable risk factors and improving postoperative management.

The present study developed a simplified predictive model to identify risk factors for postoperative UTI and assist clinicians in implementing preventive strategies. To our knowledge, this study includes one of the largest datasets evaluating pediatric UTI following unilateral pyeloplasty. The predictive model identified open surgery, left-sided UPJO, double-J stent placement, low BUN levels, and shorter height as major factors associated with postoperative UTI risk. Previous studies have reported double-J stent placement, elevated BUN levels, and shorter height as potential risk factors.[18,19] However, laterality has not previously been reported as a risk factor. In the present study, patients with left-sided UPJO were more likely to develop postoperative UTI. One possible explanation is that approx. 60% of UPJO cases are left-sided,[20] which is consistent with the findings of this study (right-sided UPJO: 33%; left-sided UPJO: 67%). Furthermore, the left ureter is generally slightly longer than the right ureter, which may contribute to greater urinary stasis and increased susceptibility to urinary reflux and infection.

In the present study, 265 (35%) children developed postoperative UTI, which is markedly higher than the previously reported rates of 7.8–10%,[21,22] but slightly lower than the 37.30% reported by Wang et al.,[19] who attributed the high incidence to the exclusion of pediatric patients receiving intraoperative antibiotics. Consistent with the findings of the present study, previous reports have suggested that indwelling double-J stents may represent a major risk factor for bacteriuria (3–10%) and may increase the risk of infection by 10–25%.[23]

According to Kabbani et al., the presence of an indwelling internal tube was identified as a significant risk factor for postoperative UTI in children after cardiac surgery, with an infection density rate of 18 per 1,000 urinary catheter-days.[24] Similarly, Wang et al. found that infants with indwelling double-J catheters were at higher risk of UTI.[7] In addition, Kitano et al. demonstrated that indwelling drainage tubes, hydronephrosis, and kidney stones were associated with an increased risk of UTI.[25] Furthermore, previous studies have suggested that prolonged indwelling drainage tube duration is associated with an elevated risk of UTI in pediatric intensive care unit (PICU) patients.[23,25] This association may result from bacterial adherence to the tube wall and subsequent biofilm formation during catheterization. Strategies such as appropriate indications for tube insertion, effective management of catheter leakage, and prompt tube removal may help reduce the risk of catheter-associated UTIs.[25]

In the present study, low BUN levels (<3.34 mmol/L) were identified as a risk factor, which is inconsistent with the findings reported by Wang et al.[19] Wang et al. suggested that BUN levels >4.3 mmol/L may also represent a risk factor, whereas the present study identified lower BUN levels (<3.4 mmol/L) as being associated with increased postoperative UTI risk. In addition, children undergoing OP were found to be at higher risk of postoperative UTI than those receiving LP. This finding is consistent with a meta-analysis of 14 studies,[18] which reported that OP (6.06%) was associated with a higher complication rate than LP (4.76%), including urinary leakage and UTI. In that analysis, urinary leakage occurred in 15 LP cases and 12 OP cases, whereas UTI developed in 9 LP cases and 14 OP cases. Similarly, Wang et al. reported that OP was associated with a higher risk of postoperative UTI compared with LP.[19]

In this investigation, LASSO regression followed by multivariable logistic regression was used to identify predictive factors associated with the risk of UTI following unilateral UPJO surgery. The integration of these statistical methods provided a comprehensive approach to model development and variable selection, enhancing the robustness and reliability of the findings. The results highlighted the importance of variables such as sex, operative time, drainage tube type, infection history, fistula history, age, BUN level, and renal cortical thickness.

Previous studies have suggested that preoperative drug susceptibility testing and prophylactic antibiotics should be considered in children at moderate or high risk of postoperative UTI.[19,26–28] However, in the present study, the incidence of infection remained high despite the use of prophylactic antibiotics. In addition, cranberry products have been reported as a potential prophylactic strategy against UTI in otherwise healthy children.[29]

Most importantly, a predictive nomogram was developed and demonstrated good discriminatory performance and calibration, as reflected by the AUC values, calibration plots, and DCA results. This nomogram may provide a useful tool for clinical risk assessment and perioperative decision-making in children undergoing unilateral UPJO surgery. Overall, the analysis of these predictive factors highlights their potential contribution to postoperative UTI risk and may support future optimization of patient management strategies.

Limitations of the study

This study has several limitations. First, this was a single-center rather than a multicenter study, and its retrospective design may have introduced bias into the results. Second, external validation was lacking. Third, there is a lack of globally representative data. As the number of patients increases, additional risk factors may be identified and more effective predictive models may be developed. Therefore, further large-scale multicenter studies are needed to validate these findings.

Conclusions

Eight risk factors, including surgical modality, laterality, drainage tube type, BUN level, and height, were identified and used to develop a predictive model for postoperative UTI in children undergoing unilateral pyeloplasty. This simplified risk-scoring model may be useful in clinical practice

by helping surgeons identify relevant preoperative risk factors, monitor children at high risk of postoperative UTI, and improve postoperative management of affected patients.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17018533. The package contains the following files:

Supplementary Table 1. Generalized variance inflation factors (GVIF) and adjusted GVIF for variables, along with Box–Tidwell and Shapiro–Wilk test results.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo at https://doi.org/10.5281/zenodo.14862979.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Hongyang Wang https://orcid.org/0000-0003-2476-4788
Chunsheng Hao https://orcid.org/0000-0002-8616-1192
Long Li https://orcid.org/0000-0003-0358-3929
Qing Sun https://orcid.org/0009-0000-4953-698X
Xiaomeng Cui https://orcid.org/0009-0002-3506-7865
Dongsheng Bai https://orcid.org/0000-0001-9619-5962
Jinqiu Song https://orcid.org/0009-0004-0498-4171

References

1. Krajewski W, Wojciechowska J, Dembowski J, Zdrojowy R, Szydełko T. Hydronephrosis in the course of ureteropelvic junction obstruction: An underestimated problem? Current opinion on pathogenesis, diagnosis and treatment. *Adv Clin Exp Med*. 2017;26(5):857–864. doi:10.17219/acem/59509
2. Nerli RB, Reddy M, Prabha V, Koura A, Patne P, Ganesh MK. Complications of laparoscopic pyeloplasty in children. *Pediatr Surg Int*. 2009;25(4):343–347. doi:10.1007/s00383-009-2341-y
3. Szavay P, Zundel S. Surgery of uretero-pelvic junction obstruction (UPJO). *Semin Pediatr Surg*. 2021;30(4):151083. doi:10.1016/j.sempedsurg.2021.151083
4. Zhang X, Li HZ, Ma X, et al. Retrospective comparison of retroperitoneal laparoscopic versus open dismembered pyeloplasty for ureteropelvic junction obstruction. *J Urol*. 2006;176(3):1077–1080. doi:10.1016/j.juro.2006.04.073
5. Crigger C, Barnard J, McClelland DJ, Ost M. Pyeloplasty. In: Gargollo PC, ed. *Minimally Invasive and Robotic-Assisted Surgery in Pediatric Urology*. Cham, Switzerland: Springer International Publishing; 2020:91–99. doi:10.1007/978-3-030-57219-8_7
6. Rai A, Hsieh A, Smith A. Contemporary diagnosis and management of pelvi-ureteric junction obstruction. *BJU Int*. 2022;130(3):285–290. doi:10.1111/bju.15689
7. Al Aaraj MS, Badreldin AM. Ureteropelvic junction obstruction. In: *StatPearls*. Treasure Island, USA: StatPearls Publishing; 2025. http://www.ncbi.nlm.nih.gov/books/NBK560740. Accessed July 29, 2025.
8. Foxman B. Urinary tract infection syndromes. *Infect Dis Clin North Am*. 2014;28(1):1–13. doi:10.1016/j.idc.2013.09.003
9. Shaikh N, Morone NE, Bost JE, Farrell MH. Prevalence of urinary tract infection in childhood: A meta-analysis. *Pediatr Infect Dis J*. 2008;27(4):302–308. doi:10.1097/INF.0b013e31815e4122
10. Nakanishi K, Okutani T, Kotani S, Kamoi Y, Kim S, Yamane M. Risk factors for cefazolin-resistant febrile urinary tract infection in children. *Pediatr Int*. 2022;64(1):e15046. doi:10.1111/ped.15046
11. Wang J, Cao Y, Zhang L, Liu G, Li C. Pathogen distribution and risk factors for urinary tract infection in infants and young children with retained double-J catheters. *J Int Med Res*. 2021;49(5):03000605211012379. doi:10.1177/03000605211012379
12. 'T Hoen LA, Bogaert G, Radmayr C, et al. Update of the EAU/ESPU guidelines on urinary tract infections in children. *J Pediatr Urol*. 2021;17(2):200–207. doi:10.1016/j.jpurol.2021.01.037
13. Gnech M, Bujons A, Radmayr C, et al. Update and summary of the EAU/ESPU paediatric guidelines on urinary tract infection in children. *J Pediatr Urol*. 2026;22(2):105481. doi:10.1016/j.jpurol.2025.06.016
14. Grasso M, Caruso RP, Phillips CK. UPJ obstruction in the adult population: Are crossing vessels significant? *Rev Urol*. 2001;3(1):42–51. PMID:16985690. PMCID:PMC1476031.
15. Kletscher BA, Segura JW, LeRoy AJ, Patterson DE. Percutaneous antegrade endopyelotomy: Review of 50 consecutive cases. *J Urol*. 1995;153(3 Pt 1):701–703. PMID:7861513.
16. Adeyoju AB, Hrouda D, Gill IS. Laparoscopic pyeloplasty: The first decade. *BJU Int*. 2004;94(3):264–267. doi:10.1111/j.1464-410X.2003.04959.x
17. Kearns J, Gundeti M. Pediatric robotic urologic surgery-2014. *J Indian Assoc Pediatr Surg*. 2014;19(3):123. doi:10.4103/0971-9261.136456
18. Chen Y, Ge XH, Yu Q, et al. Prediction model for urinary tract infection in pediatric urological surgery patients. *Front Public Health*. 2022;10:888089. doi:10.3389/fpubh.2022.888089
19. Wang H, Hao C, Bai D. Risk factors of urinary tract infection in pediatric patients with ureteropelvic junction obstruction after primary unilateral pyeloplasty. *Comput Math Methods Med*. 2022;2022:3482450. doi:10.1155/2022/3482450
20. Huang Y, Wu Y, Shan W, Zeng L, Huang L. An updated meta-analysis of laparoscopic versus open pyeloplasty for ureteropelvic junction obstruction in children. *Int J Clin Exp Med*. 2015;8(4):4922–4931. PMID:26131065. PMCID:PMC4483847.
21. He Y, Song H, Liu P, et al. Primary laparoscopic pyeloplasty in children: A single-center experience of 279 patients and analysis of possible factors affecting complications. *J Pediatr Urol*. 2020;16(3):331.e1–331.e11. doi:10.1016/j.jpurol.2020.03.028
22. Ceyhan E, Ileri F, Ceylan T, Aydin AM, Dogan HS, Tekgul S. Predictors of recurrence and complications in pediatric pyeloplasty. *Urology*. 2019;126:187–191. doi:10.1016/j.urology.2019.01.014
23. Joshi M. Urinary tract infection in the intensive care unit: A common occurrence, but with minimal clarity. *Infect Dis Clin Pract*. 2007;15(6):355–356. doi:10.1097/ipc.0b013e31815c5e82
24. Kabbani MS, Ismail SR, Fatima A, et al. Urinary tract infection in children after cardiac surgery: Incidence, causes, risk factors and outcomes in a single-center study. *J Infect Public Health*. 2016;9(5):600–610. doi:10.1016/j.jiph.2015.12.017
25. Kitano H, Shigemoto N, Koba Y, et al. Indwelling catheterization, renal stones, and hydronephrosis are risk factors for symptomatic *Staphylococcus aureus*-related urinary tract infection. *World J Urol*. 2021;39(2):511–516. doi:10.1007/s00345-020-03223-x
26. Rickard M, Dos Santos J, Keunen J, Lorenzo AJ. Prenatal hydronephrosis: Bridging pre- and postnatal management. *Prenat Diagn*. 2022;42(9):1081–1093. doi:10.1002/pd.6114
27. Renko M, Salo J, Ekstrand M, et al. Meta-analysis of the risk factors for urinary tract infection in children. *Pediatr Infect Dis J*. 2022;41(10):787–792. doi:10.1097/INF.0000000000003628
28. Dantham P, Nuvvula S, Ismail AF, Akkilagunta S, Mallineni SK. Association between passive smoking and dental caries status in children: A cross-sectional analytical study. *Dent Med Probl*. 2024;61(2):209–216. doi:10.17219/dmp/156655
29. Durham SH, Stamm PL, Eiland LS. Cranberry products for the prophylaxis of urinary tract infections in pediatric patients. *Ann Pharmacother*. 2015;49(12):1349–1356. doi:10.1177/1060028015606729

Association between rivaroxaban use and acute kidney injury in patients over 65 years: A large-scale analysis of spontaneous reporting system data

Peipei Peng[1,C,D,F], Yujie Guo[1,A–C,E,F], Yali Zhu[2,B,C,E,F]

[1] School of Nursing and Rehabilitation, Nantong University, China
[2] Department of Oncology Nursing, Nantong Tumor Hospital, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1027–1035

Address for correspondence
Yujie Guo
E-mail: guoyujievip@163.com

Funding sources
None declared

Conflict of interest
None declared

Received on March 18, 2025
Reviewed on August 3, 2025
Accepted on September 4, 2025

Published online on June 23, 2026

Abstract

Background. Given that long-term anticoagulant therapy is required in many clinical scenarios, continuous safety monitoring of direct factor Xa inhibitors is of great importance. To date, the renal safety of these agents remains a subject of debate.

Objectives. The aim of this study was to assess the association between direct factor Xa inhibitors and acute kidney injury (AKI) using data from the U.S. Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS).

Materials and methods. The study period extended from the first quarter of 2014 to the last quarter of 2024, and reports of adverse events (AEs) related to rivaroxaban and apixaban were extracted separately from FAERS. The association between direct factor Xa inhibitors and AKI was evaluated using disproportionality analysis methods.

Results. Rivaroxaban-associated AKI was more common in men, and the risk was higher among patients older than 65 years. Rivaroxaban showed a significant positive signal for AKI in patients older than 65 years, with a reported reporting odds ratio (ROR) of 3.82 (95% confidence interval (95% CI): 3.65–4.01), proportional reporting ratio (PRR) of 3.78, EBGM05 of 3.68, and IC025 of 1.88. In comparison, apixaban showed no significant risk signal, with an ROR of 0.86 (95% CI: 0.79–0.94), PRR of 0.86, EBGM05 of 0.86, and IC025 of −0.34. Acute kidney injury typically occurred at a median of 109 days after treatment initiation, with approx. 50% of cases occurring within the first 3 months. The main outcomes of AKI were hospitalization (56.06%) and death (37.67%).

Conclusions. Our findings suggest a significant association between rivaroxaban use and AKI in elderly patients. Clinical monitoring of renal function should be intensified in elderly patients receiving rivaroxaban. However, it should be emphasized that this study presents only disproportionality analysis results and cannot establish a causal relationship between rivaroxaban and AKI. Therefore, the findings should be interpreted with caution.

Key words: acute kidney injury, rivaroxaban, spontaneous reporting system, direct factor Xa inhibitors, disproportionality analysis

Cite as
Peng P, Guo Y, Zhu Y. Association between rivaroxaban use and acute kidney injury in patients over 65: A large-scale analysis of spontaneous reporting system data. *Adv Clin Exp Med*. 2026;35(6):1027–1035. doi:10.17219/acem/210295

DOI
10.17219/acem/210295

Highlights

- FAERS pharmacovigilance analysis identified a significant reporting signal for acute kidney injury (AKI) among elderly patients receiving rivaroxaban.
- Rivaroxaban-associated AKI occurred most frequently within the first 3 months of treatment and was linked to substantial rates of hospitalization and mortality.
- No significant AKI safety signal was detected for apixaban, suggesting potential differences in renal safety profiles among direct factor Xa inhibitors.
- Enhanced renal function monitoring may be warranted in older adults treated with rivaroxaban, although causal relationships cannot be established from spontaneous reporting data.

Background

Direct factor Xa inhibitors are widely used novel oral anticoagulants that exert their anticoagulant effects by directly and reversibly binding to coagulation factor Xa and inhibiting its activity without requiring a cofactor.[1] Compared with traditional anticoagulants, direct factor Xa inhibitors offer a faster onset of action, more predictable anticoagulant effects, a wider therapeutic window, no requirement for routine monitoring, and fewer interactions with food and other medications.[2] Currently, rivaroxaban and apixaban are the most commonly used agents in this class and have demonstrated particularly strong market performance. Since its approval, rivaroxaban has shown continuous global sales growth, reaching $6,568 billion in 2018. However, in recent years, its market share has been surpassed by apixaban, whose global sales reached $10,263 billion in the 1st half of 2023 alone, making it the 2nd best-selling drug worldwide.

Given that long-term anticoagulant therapy is required in many clinical settings, continuous safety monitoring of direct factor Xa inhibitors is of great importance. To date, the renal safety of these agents remains controversial. A case report described a patient receiving long-term rivaroxaban therapy who was found to have previously unrecognized glomerular disease.[3] Marcelino et al. analyzed data from regulatory centers in 134 countries and found that rivaroxaban, apixaban, and edoxaban were associated with renal adverse events (AEs) in 3.5%, 2.0%, and 1.7% of cases, respectively.[4] Among users of direct factor Xa inhibitors, acute kidney injury (AKI) is the most frequently reported renal AE, and severe cases may require renal replacement therapy.[5–10] A pharmacovigilance study based on the U.S. Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) identified a significant disproportionality signal between rivaroxaban and AKI.[11]

On the other hand, several studies have suggested potential renoprotective effects of these drugs. A meta-analysis indicated that rivaroxaban was associated with favorable renal outcomes in patients with diabetes and preexisting kidney disease, as well as in elderly populations and and Asian cohorts.[12] In patients with nonvalvular atrial fibrillation and end-stage chronic kidney disease (CKD), rivaroxaban has been associated with a significantly lower incidence of major bleeding compared with warfarin.[13] Furthermore, evidence suggests that direct factor Xa inhibitors are associated with a lower risk of AKI compared with warfarin.[14]

Against this background, further evaluation is needed to determine whether a potential safety signal exists between direct factor Xa inhibitors and AKI using real-world data. Pharmacovigilance is a key method for monitoring associations between drugs and AEs following market approval. The FAERS is a spontaneous reporting database that collects reports of suspected adverse drug reactions[15,16] and serves as an important source of information for post-marketing drug safety research.[17]

Objectives

To clearly illustrate the association between direct factor Xa inhibitors and kidney injury, this study comprehensively analyzed post-marketing kidney injury AEs associated with direct factor Xa inhibitors using real-world FAERS data. The study investigated the relationship between direct factor Xa inhibitors and kidney injury AEs and their influencing factors, compared differences in renal injury AEs between 2 direct factor Xa inhibitors, and aimed to provide a reference for their clinical use.

Materials and methods

Data source

This study was designed as an observational, retrospective, cross-sectional pharmacovigilance analysis covering the period from the 1st quarter of 2014 (Q1 2014) to the 4th quarter of 2024 (Q4 2024). Data were extracted from the FAERS, a publicly accessible large-scale spontaneous AE reporting database (https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html). As all data used in this study were anonymized and de-identified, ethical approval was not required.

Drug selection and adverse reaction definition

This study focused on 2 widely used direct factor Xa inhibitors, rivaroxaban and apixaban. To minimize potential confounding effects from concomitant medications, only reports in which rivaroxaban or apixaban was listed as the primary suspect drug were included. Reports involving co-medications were excluded. Adverse events in the FAERS database were defined using Preferred Terms (PTs) from v. 25.0 of the Medical Dictionary for Regulatory Activities (MedDRA; https://www.ich.org/page/meddra). We identified relevant AEs using the PT "acute kidney injury (AKI)" as the target keyword. Reports related to the target drugs were screened using both their generic and brand names. For each eligible report, we extracted basic information, including patient age, sex, report year, reporting country, indication, and clinical outcome. The time to onset (TTO) of AKI associated with direct factor Xa inhibitors was defined as the interval between EVENT_DT (event date) in the DEMO file and START_DT (start date of drug administration) in the THER file. Records were excluded if data were missing or inaccurate, or if the reported event occurred before the start of drug administration.

Statistical analyses

This study utilized a disproportionality analysis approach to evaluate the relative risk of AKI in patients treated with direct factor Xa inhibitors. By comparing the observed number of AE reports with the expected number, this method helps identify signals that may indicate an increased risk of adverse reactions and serves as a key tool for post-marketing drug safety surveillance. To improve the accuracy of signal detection, we employed several commonly used disproportionality measures, including reporting odds ratio (ROR),[18] proportional reporting ratio (PRR),[19] Bayesian Confidence Propagation Neural Network (BCPNN),[20] and Multi-item Gamma Poisson Shrinker (MGPS).[21] A signal was considered positive only if the predefined threshold criteria for all 4 algorithms were met. The specific mathematical formulas and evaluation criteria are detailed in Table 1. All disproportionality indicators were calculated based on 2 × 2 contingency tables, as presented in Table 2.

To assess the robustness of our findings, ibuprofen was selected as a positive control and insulin glargine as a negative control. Additionally, subgroup analyses were performed based on sex and age groups to enhance the interpretability and reliability of the results.

All data processing and statistical analyses were conducted using Microsoft Excel 2019 (Microsoft Corp., Redmond, USA) and R v. 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria). The primary R packages used for data cleaning, analysis, and visualization included "desc", "tidyverse", "table1", "openxlsx", "data.table", "dplyr", and "ggplot2".

Table 1. Summary of the major algorithms used for signal detection

Algorithms	Equation	Criteria
ROR	$ROR = (a/b)/(c/d)$	lower 95% CI > 1, n ≥ 3
	$95\%\ CI = e^{\ln(ROR) \pm 1.96(1/a+1/b+1/c+1/d)^{0.5}}$	
PRR	$PRR = (a/(a+c))/(b/(b+d))$	PRR ≥ 2, χ^2 ≥ 4, n ≥ 3
	$X^2 = \sum((O-E)^2/E)$; (O = a, E = (a+ b)(a+c)/(a+b+c+d))	
BCPNN	$a_{exp} = ((a+b)*(a+c))/(a+b+c+d)$	$IC_{025} > 0$
	$IC = \log_2((a+0.5)/(a_{exp}+0.5))$	
	$IC_{025} = IC\text{-}3.3*(a+0.5)^{-0.5} - 2*(a+0.5)^{-1.5}$	
MGPS	$EBGM = a(a+b+c+d)/((a+c)(a+b))$	EBGM05 > 2, n > 3
	$EBGM05 = e^{\ln(EBGM)-1.64(1/a+1/b+1/c+1/d)^{0.5}}$	

a – report count of the target adverse event for the target drug; b – report count of other adverse events for the target drug; c – report count of the target adverse event for all other drugs; d – report count of other adverse events for all other drugs; ROR – reporting odds ratio; PRR – proportional reporting ratio; BCPNN – Bayesian Confidence Propagation Neural Network; MGPS – Multi-item Gamma Poisson Shrinker; IC – information component; EBGM – empirical Bayes geometric mean; 95% CI – 95% confidence interval.

Table 2. A 2 × 2 contingency table used for disproportionality analysis

Number of cases	Direct factor Xa inhibitors	Other drugs
Acute kidney injury	a	c
Non-acute kidney injury	b	d

a – report count of the target adverse event for the target drug; b – report count of other adverse events for the target drug; c – report count of the target adverse event for all other drugs; d – report count of other adverse events for all other drugs.

Results

Basic characteristics

A total of 108,819 AE reports related to rivaroxaban and 135,127 reports related to apixaban were retrieved from the FAERS database between Q1 2014 and Q4 2024. Among them, 2,479 reports of AKI were associated with rivaroxaban (2.28%), while 601 AKI cases were associated with apixaban (0.44%). The annual number of AE reports related to direct factor Xa inhibitors over the past decade is shown in Fig. 1.

Descriptive analyses of the clinical characteristics of the reported cases are summarized in Table 3. In both drug groups, AKI cases were predominantly reported in male patients and primarily occurred in individuals aged 65 years and older. Notably, AKI reports were rare among individuals younger than 18 years, likely reflecting the demographic profile of the treated population. The majority of AKI-related reports originated from the USA and France. The most frequently reported indications for direct factor Xa inhibitors included atrial fibrillation, cerebrovascular accident prophylaxis, thrombosis prophylaxis, pulmonary embolism, and hypertension.

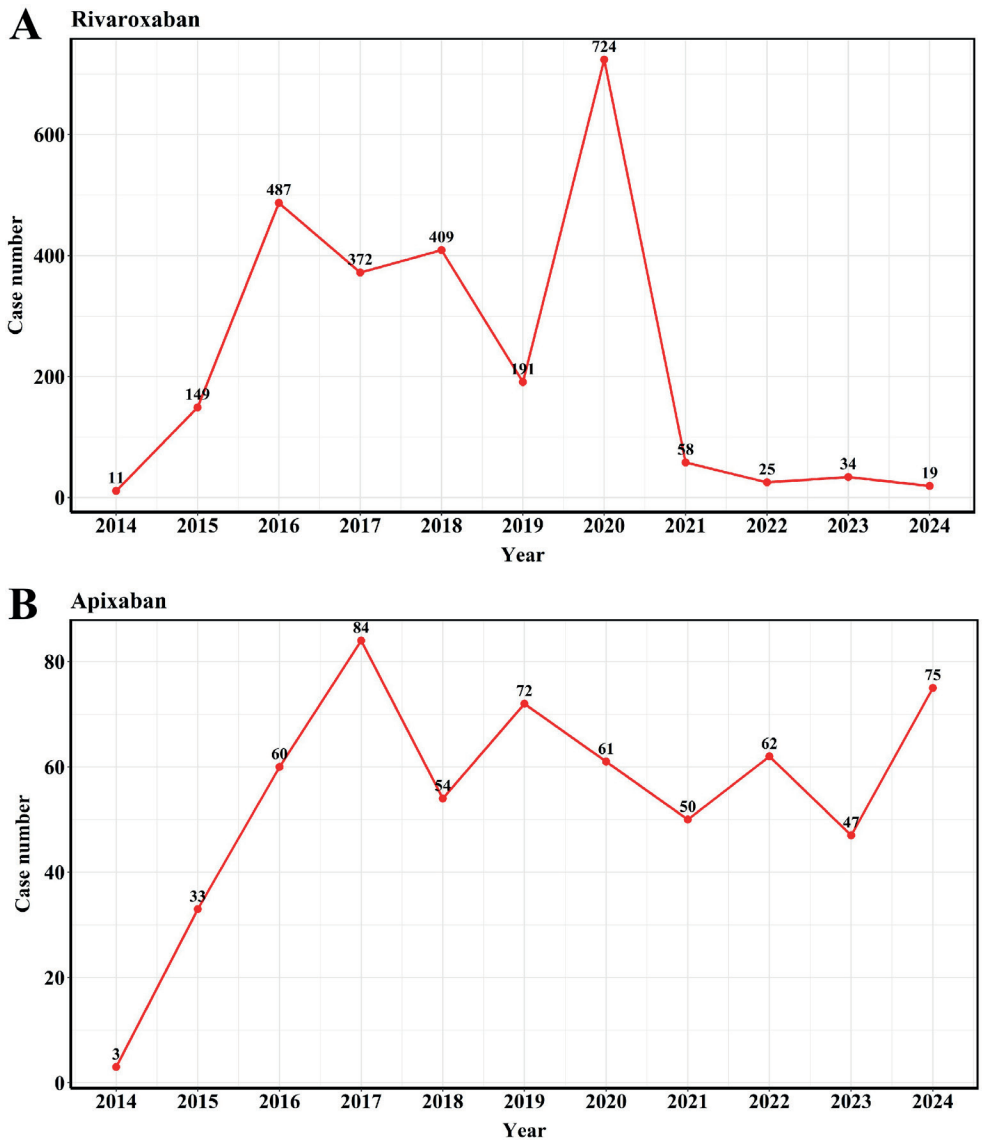


Fig. 1. Annual number of adverse event reports related to direct factor Xa inhibitors over the past decade

Disproportionality analysis in elderly patients aged over 65 years

Since most AKI reports occurred in elderly patients (≥65 years), disproportionality analysis was limited to this subgroup, using the entire FAERS database as the reference. To validate the analytical performance of FAERS, ibuprofen and insulin glargine were used as positive and negative controls, respectively. As shown in Table 4, all 4 disproportionality methods indicated a strong association between ibuprofen and AKI (ROR = 5.75, 95% confidence interval (95% CI): 5.31–6.22; PRR = 5.63, χ^2 = 2,377.41; EBGM = 5.58, EBGM05 = 5.22; IC = 2.47, IC025 = 2.34), whereas no significant association was observed for insulin glargine (ROR = 0.11, 95% CI: 0.08–0.16; PRR = 0.11; EBGM = 0.12, EBGM05 = 0.09; IC = −3.10, IC025 = −3.68). Subsequent disproportionality analysis using 4 methods (ROR, PRR, EBGM, and IC) revealed a significant signal for the association between rivaroxaban and AKI in patients aged ≥65 years, while no such signal was observed for apixaban (Table 4). Rivaroxaban exhibited a high ROR (3.82, 95% CI: 3.65–4.01), whereas apixaban had a lower ROR (0.86, 95% CI: 0.79–0.94). The PRR results were consistent with the ROR findings: the PRR for rivaroxaban was 3.78 (χ^2 = 3543.98), whereas that for apixaban was 0.78 (χ^2 = 10.73). Further validation using EBGM and IC also confirmed these findings: rivaroxaban showed EBGM = 3.68 (EBGM05 = 3.54) and IC = 1.88 (IC025 = 1.80), both exceeding the signal thresholds. In contrast, apixaban had EBGM = 0.86 (EBGM05 = 0.80) and IC = −0.21 (IC025 = −0.36), indicating no signal. Based on these results, subsequent analyses focused on rivaroxaban, which showed a significant association with AKI.

Comparative analysis with other antithrombotic agents

In addition to direct factor Xa inhibitors, commonly used oral antithrombotic agents include warfarin and aspirin. Figure 2 displays the variation in AKI signals for rivaroxaban and apixaban using warfarin or aspirin as reference

Table 3. Characteristics of acute kidney injury events associated with direct factor Xa inhibitors reported in the U.S. Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) database, 2014–2024

Characteristics		Rivaroxaban, n (%)	Apixaban, n (%)
All adverse events		108,819	135,127
Acute kidney injury		2,479 (2.28%)	601 (0.44%)
Gender	male	1,379 (1.27%)	320 (0.24%)
	female	1,066 (0.98%)	255 (0.19%)
	not reported	34 (0.03%)	26 (0.02%)
Age groups	≤18	2 (0.00%)	0 (0.00%)
	18 to <65	598 (0.55%)	51 (0.04%)
	65 to ≤75	829 (0.76%)	152 (0.11%)
	>75	960 (0.88%)	340 (0.25%)
	not reported	90 (0.08%)	58 (0.04%)
Reported occupation	consumer	1,809 (1.66%)	120 (0.09%)
	physician	238 (0.22%)	208 (0.15%)
	health professional	209 (0.19%)	89 (0.07%)
	other health professional	98 (0.09%)	50 (0.04%)
	pharmacist	98 (0.09%)	127 (0.09%)
	not reported	4 (0.00%)	3 (0.00%)
Reported country	USA	2,106 (1.94%)	184 (0.14%)
	France	88 (0.08%)	128 (0.09%)
	Germany	51 (0.05%)	52 (0.04%)
	UK	50 (0.05%)	57 (0.04%)
	Italy	25 (0.02%)	33 (0.02%)
Indication	atrial fibrillation	695 (0.64%)	90 (0.07%)
	cerebrovascular accident prophylaxis	577 (0.53%)	290 (0.21%)
	thrombosis prophylaxis	341 (0.31%)	22 (0.02%)
	deep vein thrombosis	336 (0.31%)	17 (0.01%)
	pulmonary embolism	154 (0.14%)	21 (0.02%)

Table 4. Disproportionality analysis of acute kidney injury (AKI) events associated with direct factor Xa inhibitors in elderly patients aged over 65 years

Drug	ROR (95% CI)	PRR (χ^2)	EBGM (EBGM05)	IC (IC_{025})
Ibuprofen (positive control)	5.75 (5.31–6.22)	5.63 (2,377.41)	5.58 (5.22)	2.48 (2.36)
Insulin glargine (negative control)	0.11 (0.08–0.16)	0.11 (226.54)	0.12 (0.09)	−3.12 (−3.61)
Rivaroxaban	3.82 (3.65–4.01)	3.78 (3,543.98)	3.68 (3.54)	1.88 (1.81)
Apixaban	0.86 (0.79–0.94)	0.86 (10.73)	0.86 (0.80)	−0.21 (−0.34)

Values in bold indicate a positive signal. ROR – reporting odds ratio; PRR – proportional reporting ratio; BCPNN – Bayesian Confidence Propagation Neural Network; MGPS – Multi-item Gamma Poisson Shrinker; IC – information component; EBGM – empirical Bayes geometric mean; 95% CI – 95% confidence interval..

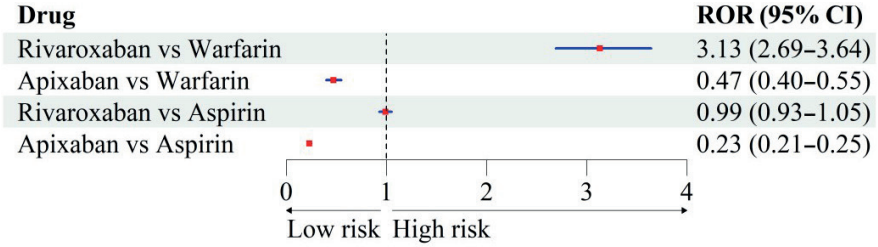


Fig. 2. Changes in acute kidney injury (AKI) signals for rivaroxaban and apixaban when warfarin or aspirin was used as the reference drug

ROR – reporting odds ratio; 95% CI – 95% confidence interval.

drugs. Compared with warfarin, rivaroxaban showed a significant association with AKI (ROR = 3.13, 95% CI: 2.69–3.64), whereas apixaban did not (ROR = 0.47, 95% CI: 0.40–0.55). When compared with aspirin, neither rivaroxaban (ROR = 0.99, 95% CI: 0.93–1.05) nor apixaban (ROR = 0.23, 95% CI: 0.21–0.25) showed a significant signal.

Table 5. Rivaroxaban-associated AKI according to clinical characteristics in patients aged over 65 years

Characteristics		Patients (total)	Number of AKI cases	Proportion of AKI (%)	ROR (95% CI)
Sex	female	47,878	1,066	2.23%	6.05 (5.69–6.43)
	male	48,145	1,379	2.86%	4.12 (3.90–4.35)
Age [years]	65–75	21,823	829	3.80%	4.55 (4.24–4.88)
	>75	26,757	960	3.59%	3.11 (2.91–3.31)

AKI – acute kidney injury; ROR – reporting odds ratio; 95% CI – 95% confidence interval.

Subgroup analysis in patients aged over 65 years

Given the observed association between rivaroxaban and AKI in older adults, a stratified analysis by sex and age was conducted to control for potential confounders (Table 5). Reporting odds ratio values for rivaroxaban remained above the signal threshold across all subgroups. Notably, the risk was more pronounced in elderly females, with an ROR of 6.05 (95% CI: 5.69–6.43), indicating higher susceptibility in this population. It should be noted that these findings warrant confirmation in larger, high-quality prospective studies.

Time to onset

The median time to AKI onset following rivaroxaban administration was 109 days (interquartile range (IQR): 22–340 days). Approximately 50% of AKI events occurred within the first 3 months of treatment. Interestingly, about 10% of AKI cases emerged within the 1st week of drug exposure (Fig. 3).

Prognosis of rivaroxaban-associated AKI in patients over 65 years

To assess the clinical outcomes of rivaroxaban-associated AKI, hospitalization and fatality rates were analyzed using the FAERS database (Table 6). The results indicated a hospitalization rate of 56.06%, a fatality rate of 37.67%, and a life-threatening event rate of 4.02%, all of which were significantly higher than those observed in patients without AKI. These findings highlight the substantial impact of AKI on clinical outcomes in this population.

Table 6. Outcomes of AKI events in elderly patients aged over 65 years

Outcome	Rivaroxaban	
	AKI (n = 1,789)	non-AKI (n = 46,791)
Death	674 (37.67%)	8,866 (18.95%)
Hospitalization	1,003 (56.06%)	22,264 (47.58%)
Life-threatening events	72 (4.02%)	2,089 (4.47%)
Other serious events	34 (1.90%)	8,800 (18.81%)

AKI – acute kidney injury.

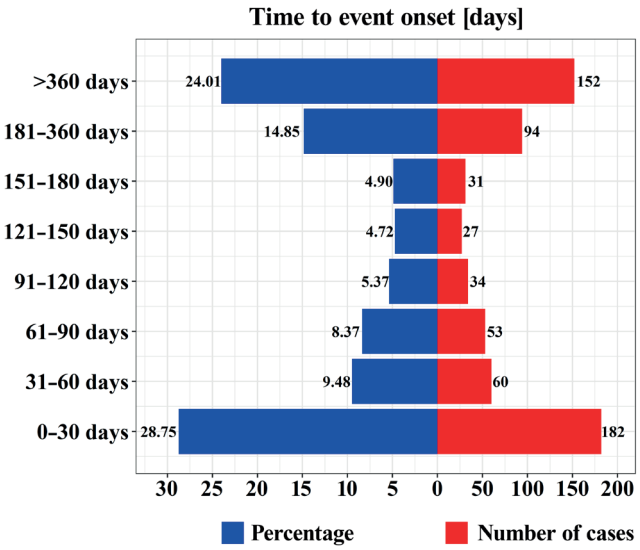


Fig. 3. Time to onset of acute kidney injury (AKI) following rivaroxaban administration

Discussion

Currently, post-marketing safety data regarding direct factor Xa inhibitors and their use in renal disorders remain limited. The FAERS database, as one of the major global resources for detecting signals of rare AEs,[22] provides invaluable data for pharmacovigilance and pharmacoepidemiological research and is freely accessible to researchers. In this study, we systematically investigated the potential renal adverse effects associated with direct factor Xa inhibitors using the FAERS surveillance system.

From spontaneous reports submitted between 2014 and 2024, we identified 2,479 AKI cases associated with rivaroxaban and 601 associated with apixaban. Although FAERS data cannot be used to estimate the true incidence rate, the proportion of AKI among all rivaroxaban-related AEs was 2.28% (2,479/108,819), suggesting that this AE may not be as rare as described in the product label. The main strength of this study lies in being the first to systematically evaluate the potential association between direct factor Xa inhibitors, particularly rivaroxaban, and AKI. All 4 disproportionality analysis methods indicated a significant signal for rivaroxaban-associated AKI: ROR = 3.82 (95% CI: 3.65–4.01), PRR = 3.78 (χ^2 = 3,543.98), EBGM = 3.68 (EBGM05 = 3.54), and IC = 1.88 (IC025 = 1.81). In contrast, no statistically significant association was observed for apixaban.

With the growing number of patients with cardiovascular disease (CVD), the use of oral anticoagulants has become increasingly common. As early as 2017, case reports suggested that direct factor Xa inhibitors may induce nephrotoxicity, supported by renal biopsy findings in a patient treated with apixaban.[23,24] Subsequent reports by Galloway et al.[25] and Iwafuchi et al.[26] also described apixaban-associated AKI cases. However, our analysis indicated a significant signal only between rivaroxaban and AKI, with no similar findings for apixaban. Although both agents belong to the same class, emerging evidence suggests potential differences in renal safety profiles. Several comparative studies and meta-analyses have shown that apixaban may pose a lower risk of AKI than rivaroxaban in elderly individuals and those with CKD or diabetes.[27] Apixaban has not been statistically linked to adverse renal outcomes in most patients with atrial fibrillation.[27] This disparity may arise from differences in pharmacokinetics, renal clearance, and dosing regimens. For example, apixaban is less reliant on renal elimination than rivaroxaban, which may contribute to its more favorable renal safety profile. Additionally, proposed mechanisms of AKI induced by direct factor Xa inhibitors include glomerular hemorrhage, tubular obstruction by red blood cell casts, and free hemoglobin release.[28] Users of apixaban had a lower risk of bleeding compared with those receiving rivaroxaban.[29]

We further analyzed the effects of sex and age on AKI risk. Rivaroxaban-associated AKI was more frequently reported in men and in patients aged ≥65 years. All 4 disproportionality methods confirmed a clear signal for rivaroxaban in the elderly subgroup. Epidemiological data indicate that individuals aged ≥65 years are at greater risk of cardiovascular events[30] and experience higher rates of mortality[31] and post-hospitalization disability,[32] possibly due to declining organ function and metabolic capacity.[33] Globally, the age-standardized mortality rate for CVD in men (280.8 per 100,000) is significantly higher than in women (204.0 per 100,000), particularly in high-income regions and across Central/Eastern Europe, Central Asia, and South Asia.[34] Thus, male patients are more likely to be prescribed rivaroxaban, which may explain the higher number of AE reports in this group. Clinicians should exercise greater caution when prescribing rivaroxaban to male and elderly patients. With respect to geographical distribution, the majority of rivaroxaban- and apixaban-related AEs were reported in the USA and France, suggesting international differences in reporting behavior, potentially influenced by prescribing habits, regulatory practices, and awareness of the FAERS database. Time-to-onset analysis revealed that nearly 50% of AKI events occurred within 3 months of rivaroxaban initiation, with approx. 10% occurring within the 1st week. Although current clinical guidelines do not recommend routine renal monitoring for all rivaroxaban users,[35,36] our findings underscore the importance of early renal function surveillance to mitigate potential risks.

In terms of clinical outcomes, among the 1,789 rivaroxaban-associated AKI cases, 1,003 (56.06%) required hospitalization and 674 (37.67%) resulted in death, underscoring the severity of these events. In a study by Chen et al., 35% of 28 patients with direct oral anticoagulants (DOAC)-related AKI required acute dialysis, and 60% did not recover baseline renal function.[37] Trujillo et al. similarly reported that 28% of patients receiving DOACs required long-term dialysis.[38] These findings suggest that rivaroxaban-associated AKI may synergize with complications such as bleeding or infection, complicating treatment and increasing the healthcare burden.[35] Nevertheless, this study cannot establish a causal relationship between rivaroxaban use and hospitalization or death, as these outcomes may also result from underlying disease progression or other clinical events. Prospective studies are warranted to verify these associations.

Importantly, although we observed disproportionality in the reporting of rivaroxaban-associated AKI, these findings do not imply causality and should not be overinterpreted. Numerous studies have demonstrated the therapeutic benefits of rivaroxaban in thrombotic disorders. For example, compared with aspirin, rivaroxaban combination therapy reduced all-cause and cardiovascular mortality in patients with chronic coronary or peripheral artery disease.[36] Moreover, both rivaroxaban (hazard ratio (HR) = 0.66, 95% CI: 0.55–0.77) and apixaban (HR = 0.73, 95% CI: 0.59–0.87) have been shown to reduce the risk of renal impairment.[39] In patients with moderate renal impairment and no increased bleeding risk, the benefit of rivaroxaban plus aspirin was preserved compared with aspirin alone.[40] These findings are consistent with our results showing that neither rivaroxaban nor apixaban exhibited a significant association with AKI when aspirin was used as the comparator. Furthermore, rivaroxaban may provide greater benefits in patients with CKD or diabetes. In patients with atrial fibrillation, rivaroxaban has been associated with better renal outcomes than vitamin K antagonists (VKAs).[12,27] Among patients with nonvalvular atrial fibrillation and diabetes, rivaroxaban was associated with a lower risk of renal AEs than warfarin.[41] In patients with both atrial fibrillation and advanced CKD, rivaroxaban use was associated with fewer renal AEs compared with VKAs.[42] In elderly patients, rivaroxaban tended to improve primary safety outcomes regardless of renal function.[43] However, our study identified a significant AKI signal for rivaroxaban, but not for apixaban, when compared with warfarin, which may differ from previous findings due to differences between real-world and case-control study designs.

Limitations of the study

This study has several strengths, including a large sample size, broad coverage of real-world data, and a systematic evaluation of the clinical characteristics and outcomes associated with rivaroxaban-related AKI. Nevertheless, it also

has several limitations. First, FAERS data may be incomplete, and AKI diagnoses are often based on spontaneous reports without standardized criteria. Second, the lack of control for confounding factors (e.g., pre-existing CKD) limits causal inference. Third, missing baseline characteristics, exposure information, and renal function data hinder the estimation of incidence and attribution of causality.

Conclusions

Using real-world data from FAERS, this study identified a potential safety signal linking rivaroxaban to AKI in elderly patients (≥65 years). Given that most AKI cases occurred within 3 months of treatment initiation and were associated with high hospitalization and mortality rates, clinicians should remain vigilant and conduct regular renal monitoring, especially during the early phase of therapy. Our findings provide supplementary evidence for the rational use of rivaroxaban and highlight the need to monitor renal function in older adults. It is essential to emphasize that the observed disproportionality does not confirm a causal relationship and should not lead to overinterpretation of the results.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo at https://doi.org/10.5281/zenodo.15209285.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Peipei Peng https://orcid.org/0009-0006-8305-9770
Yujie Guo https://orcid.org/0009-0005-1003-6925
Yali Zhu https://orcid.org/0009-0004-2746-8516

References

1. Zheng W, Dai X, Xu B, Tian W, Shi J. Discovery and development of Factor Xa inhibitors (2015–2022). *Front Pharmacol*. 2023;14:1105880. doi:10.3389/fphar.2023.1105880
2. Ruff CT, Giugliano RP, Braunwald E, et al. Comparison of the efficacy and safety of new oral anticoagulants with warfarin in patients with atrial fibrillation: A meta-analysis of randomised trials. *Lancet*. 2014;383(9921):955–962. doi:10.1016/S0140-6736(13)62343-0
3. Marshall L, Ahmed N, De Zoysa JR, Hayat A. Rivaroxaban-related nephropathy: An under-reported cause of acute kidney injury. *Int Med J*. 2025;55(5):864–865. doi:10.1111/imj.70080
4. Marcelino G, Hemett OM, Descombes E. Acute renal failure in a patient with rivaroxaban-induced hypersensitivity syndrome: A case report with a review of the literature and of pharmacovigilance registries. *Case Rep Nephrol*. 2020;2020:6940183. doi:10.1155/2020/6940183
5. Belčič Mikič T, Kojc N, Frelih M, Aleš-Rigler A, Večerić-Haler Ž. Management of anticoagulant-related nephropathy: A single center experience. *J Clin Med*. 2021;10(4):796. doi:10.3390/jcm10040796
6. Dantec A, Virot JS, Filancia A, Angonin R, Nefti H. Acute tubulo-interstitial nephritis due to rivaroxaban (Xarelto) [in French]. *Presse Med*. 2017;46(5):541–542. doi:10.1016/j.lpm.2017.03.019
7. Monahan RC, Suttorp MM, Gabreëls BATF. A case of rivaroxaban-associated acute tubulointerstitial nephritis. *Neth J Med*. 2017;75(4):169–171. PMID:28522776.
8. Oliveira M, Lima C, Góis M, Viana H, Carvalho F, Lemos S. Rivaroxaban-related nephropathy. *Port J Nephrol Hypert*. 2017;31(3):212–216. https://cdn02.spnefro.pt/pjnh/65/n3_2017_pjnh_11.pdf.
9. Chung EY, Chen J, Roxburgh S. A case report of Henoch–Schönlein purpura and IgA nephropathy associated with rivaroxaban. *Nephrology*. 2018;23(3):289–290. doi:10.1111/nep.13058
10. Fujino Y, Takahashi C, Mitsumoto K, Uzu T. Rivaroxaban-related acute kidney injury in a patient with IgA vasculitis. *BMJ Case Rep*. 2019;12(1):e227756. doi:10.1136/bcr-2018-227756
11. Bai H, Wang J, Nie X, et al. Associations of acute kidney injury with oral anticoagulants: A disproportionality analysis of the Food and Drug Administration Adverse Event Reporting System (FAERS) database [published online as ahead of print on February 2, 2025]. *Exp Opin Drug Saf*. 2025. doi:10.1080/14740338.2025.2461203
12. Dinh PP, Ho THQ, Pham HM, et al. Evaluating renal benefits of rivaroxaban versus vitamin K antagonists in atrial fibrillation: A systematic review and meta-analysis of real-world evidence. *Eur Cardiol*. 2024;19:e05. doi:10.15420/ecr.2024.07
13. Coleman CI, Kreutz R, Sood NA, et al. Rivaroxaban versus warfarin in patients with nonvalvular atrial fibrillation and severe kidney disease or undergoing hemodialysis. *Am J Med*. 2019;132(9):1078–1083. doi:10.1016/j.amjmed.2019.04.013
14. Derebail VK, Rheault MN, Kerlin BA. Role of direct oral anticoagulants in patients with kidney disease. *Kidney Int*. 2020;97(4):664–675. doi:10.1016/j.kint.2019.11.027
15. Patek TM, Teng C, Kennedy KE, Alvarez CA, Frei CR. Comparing acute kidney injury reports among antibiotics: A pharmacovigilance study of the FDA Adverse Event Reporting System (FAERS). *Drug Saf*. 2020;43(1):17–22. doi:10.1007/s40264-019-00873-8
16. Zhu J, He Z, Liang D, Yu X, Qiu K, Wu J. Pulmonary tuberculosis associated with immune checkpoint inhibitors: A pharmacovigilance study. *Thorax*. 2022;77(7):721–723. doi:10.1136/thoraxjnl-2021-217575
17. Zhu J, Wu J, Chen P, et al. Acute kidney injury associated with immune checkpoint inhibitors: A pharmacovigilance study. *Int Immunopharmacol*. 2022;113:109350. doi:10.1016/j.intimp.2022.109350
18. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. *Pharmacoepidemiol Drug Saf*. 2004;13(8):519–523. doi:10.1002/pds.1001
19. Evans SJW, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. *Pharmacoepidemiol Drug Saf*. 2001;10(6):483–486. doi:10.1002/pds.677
20. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol*. 1998;54(4):315–321. doi:10.1007/s002280050466
21. Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's Spontaneous Reports Database. *Drug Saf*. 2002;25(6):381–392. doi:10.2165/00002018-200225060-00001
22. Moeckel GW, Luciano RL, Brewster UC. Warfarin-related nephropathy in a patient with mild IgA nephropathy on dabigatran and aspirin. *Clin Kidney J*. 2013;6(5):507–509. doi:10.1093/ckj/sft076
23. Ryan M, Ware K, Qamri Z, et al. Warfarin-related nephropathy is the tip of the iceberg: Direct thrombin inhibitor dabigatran induces glomerular hemorrhage with acute kidney injury in rats. *Nephrol Dial Transplant*. 2014;29(12):2228–2234. doi:10.1093/ndt/gft380
24. Brodsky SV, Mhaskar NS, Thiruveedi S, et al. Acute kidney injury aggravated by treatment initiation with apixaban: Another twist of anticoagulant-related nephropathy. *Kidney Res Clin Pract*. 2017;36(4):387–392. doi:10.23876/j.krcp.2017.36.4.387
25. Galloway M, Sim JJ, Slater A, Bray C, Bishev D, Walker P. An overlap of anticoagulant-related and IgA nephropathy: A case report. *Glomerular Dis*. 2024;4(1):167–171. doi:10.1159/000541116

26. Iwafuchi Y, Ito Y, Imai N, Oyama Y, Narita I. Dabigatran-related nephropathy complicated by tubulointerstitial nephritis in a patient with a normal renal function and undiagnosed IgA nephropathy. *Intern Med*. 2024;63(11):1615–1621. doi:10.2169/internalmedicine.2628-23
27. Yao X, Tangri N, Gersh BJ, et al. Renal outcomes in anticoagulated patients with atrial fibrillation. *J Am Coll Cardiol*. 2017;70(21):2621–2632. doi:10.1016/j.jacc.2017.09.1087
28. Brodsky SV. Anticoagulants and acute kidney injury: Clinical and pathology considerations. *Kidney Res Clin Pract*. 2014;33(4):174–180. doi:10.1016/j.krcp.2014.11.001
29. Dawwas GK, Leonard CE, Lewis JD, Cuker A. Risk for recurrent venous thromboembolism and bleeding with apixaban compared with rivaroxaban: An analysis of real-world data. *Ann Intern Med*. 2022; 175(1):20–28. doi:10.7326/M21-0717
30. Rodgers JL, Jones J, Bolleddu SI, et al. Cardiovascular risks associated with gender and aging. *J Cardiovasc Dev Dis*. 2019;6(2):19. doi:10.3390/jcdd6020019
31. Kulmala J, Nykänen I, Hartikainen S. Frailty as a predictor of all-cause mortality in older men and women: Frailty and mortality. *Geriatr Gerontol Int*. 2014;14(4):899–905. doi:10.1111/ggi.12190
32. Buurman BM, Hoogerduijn JG, De Haan RJ, et al. Geriatric conditions in acutely hospitalized older patients: Prevalence and one-year survival and functional decline. *PLoS One*. 2011;6(11):e26951. doi:10.1371/journal.pone.0026951
33. Kim IH, Kisseleva T, Brenner DA. Aging and liver disease. *Curr Opin Gastroenterol*. 2015;31(3):184–191. doi:10.1097/MOG.0000000000000176
34. Institute for Health Metrics and Evaluation (IHME). GBD Compare Data Visualization. Seattle, USA: Institute for Health Metrics and Evaluation (IHME); 2023. https://vizhub.healthdata.org/gbd-compare. Accessed March 18, 2023.
35. Weber J, Olyaei A, Shatzel J. The efficacy and safety of direct oral anticoagulants in patients with chronic renal insufficiency: A review of the literature. *Eur J Haematol*. 2019;102(4):312–318. doi:10.1111/ejh.13208
36. Eikelboom JW, Bhatt DL, Fox KAA, et al. Mortality benefit of rivaroxaban plus aspirin in patients with chronic coronary or peripheral artery disease. *J Am Coll Cardiol*. 2021;78(1):14–23. doi:10.1016/j.jacc.2021.04.083
37. Chen S, Liao D, Yang M, Wang S. Anticoagulant-related nephropathy induced by direct-acting oral anticoagulants: Clinical characteristics, treatments and outcomes. *Thromb Res*. 2023;222:20–23. doi:10.1016/j.thromres.2022.12.002
38. Trujillo H, Sandino J, Cavero T, et al. IgA nephropathy is the most common underlying disease in patients with anticoagulant-related nephropathy. *Kidney Int Rep*. 2022;7(4):831–840. doi:10.1016/j.ekir.2022.01.1048
39. Zhang C, Gu ZC, Ding Z, et al. Decreased risk of renal impairment in atrial fibrillation patients receiving non-vitamin K antagonist oral anticoagulants: A pooled analysis of randomized controlled trials and real-world studies. *Thromb Res*. 2019;174:16–23. doi:10.1016/j.thromres.2018.12.010
40. Fox KAA, Eikelboom JW, Shestakovska O, Connolly SJ, Metsarinne KP, Yusuf S. Rivaroxaban plus aspirin in patients with vascular disease and renal dysfunction. *J Am Coll Cardiol*. 2019;73(18):2243–2250. doi:10.1016/j.jacc.2019.02.048
41. Hernandez AV, Bradley G, Khan M, et al. Rivaroxaban vs warfarin and renal outcomes in non-valvular atrial fibrillation patients with diabetes. *Eur Heart J Qual Care Clin Outcomes*. 2020;6(4):301–307. doi:10.1093/ehjqcco/qcz047
42. Kreutz R, Deray G, Floege J, et al. Rivaroxaban vs vitamin K antagonist in patients with atrial fibrillation and advanced chronic kidney disease. *JACC Adv*. 2024;3(2):100813. doi:10.1016/j.jacadv.2023.100813
43. Hori M, Matsumoto M, Tanahashi N, et al. Rivaroxaban vs warfarin in Japanese patients with non-valvular atrial fibrillation in relation to age. *Circ J*. 2014;78(6):1349–1356. doi:10.1253/circj.CJ-13-1324

Mendelian randomization analysis of metabolic blood biomarkers and gastrointestinal cancer risk

Xiufeng Huang[A,D,F], Jianxin Chen[A,D–F], Feiteng Gu[A,C,F], Xiaodan Zheng[A,C,F]

Second Department of Gastrointestinal Surgery, The Affiliated Hospital of Putian University, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1037–1045

Address for correspondence
Jianxin Chen
E-mail: lhxlptxy@sina.com

Funding sources
None declared

Conflict of interest
None declared

Received on May 28, 2025
Reviewed on July 25, 2025
Accepted on August 27, 2025

Published online on June 25, 2026

Abstract

Background. Gastrointestinal (GI) cancers remain among the most lethal malignancies worldwide, highlighting the need for novel insights into modifiable risk factors. Blood-based metabolic biomarkers – routinely measured in clinical settings – may play a role in cancer etiology, but their causal impact remains unclear.

Objectives. This study aimed to explore the causal relationship between metabolic blood biomarkers and the risk of GI cancers using Mendelian randomization (MR).

Materials and methods. We conducted an MR analysis using large-scale genome-wide association study (GWAS) data to evaluate the causal effects of metabolic blood traits (e.g., low-density lipoprotein (LDL) cholesterol, creatinine, uric acid, total protein, and total cholesterol) on the risk of 4 major GI cancers: gastric, colorectal, pancreatic, and gallbladder cancer. Statistical robustness was evaluated using the MR-PRESSO test and Cochran's Q test.

Results. Mendelian randomization analysis revealed several significant associations. Low-density lipoprotein cholesterol was inversely associated with gastric cancer risk (odds ratio (OR) = 0.815, 95% confidence interval (95% CI): 0.703–0.946, $p < 0.01$), while total cholesterol also showed a protective effect (OR = 0.790, 95% CI: 0.693–0.900, $p < 0.01$). Serum creatinine levels were strongly associated with a reduced risk of gallbladder cancer (OR = 0.037, 95% CI: 0.005–0.287, $p < 0.01$). Higher serum total protein (OR = 0.855, 95% CI: 0.777–0.940, $p < 0.01$) and uric acid levels (OR = 0.855, 95% CI: 0.769–0.951, $p < 0.01$) were associated with a reduced risk of colorectal cancer (CRC). Serum uric acid levels were also associated with a reduced risk of gastric cancer (OR = 0.822, 95% CI: 0.719–0.941, $p < 0.01$), but with an increased risk of pancreatic cancer (OR = 1.246, 95% CI: 1.01–1.539, $p = 0.040$).

Conclusions. Our findings provide causal evidence linking common metabolic blood biomarkers to site-specific GI cancer risks. These results may inform biomarker-based risk stratification and preventive strategies in oncology and public health.

Key words: cancer risk, genetic variants, gastrointestinal cancer, Mendelian randomization, metabolic blood biomarkers

Cite as
Huang X, Chen J, Gu F, Zheng X. Mendelian randomization analysis of metabolic blood biomarkers and gastrointestinal cancer risk. *Adv Clin Exp Med*. 2026;35(6):1037–1045. doi:10.17219/acem/209956

DOI
10.17219/acem/209956

Highlights

- Low-density lipoprotein (LDL) and total cholesterol showed protective effects against gastric cancer (odds ratio (OR) = 0.815 and OR = 0.790, respectively; $p < 0.01$), suggesting a lipid-related mechanism in gastric carcinogenesis.
- Serum creatinine significantly reduced the risk of gallbladder cancer (OR = 0.037, $p < 0.01$), highlighting a potential novel metabolic biomarker for risk assessment.
- Higher serum total protein and uric acid levels were protective against colorectal cancer (CRC; OR = 0.855 for both, $p < 0.01$), underscoring the role of metabolic balance in CRC prevention.
- Uric acid displayed dual effects, being protective against gastric cancer but associated with increased pancreatic cancer risk (OR = 0.822 vs OR = 1.246), indicating tumor site-specific metabolic pathways.

Background

Gastrointestinal (GI) cancers, including gastric, colorectal, pancreatic, and gallbladder cancers, are among the leading causes of cancer-related mortality worldwide.[1,2] According to recent global cancer statistics, gastric cancer ranks 5th in incidence and 4th in mortality, while colorectal cancer (CRC) ranks 3rd and 2nd, respectively.[3] These cancers have complex etiologies involving genetic, environmental, and lifestyle factors.[4,5] Gastric cancer, for instance, is influenced by *Helicobacter pylori* infection, dietary factors, and genetic predisposition.[6] Colorectal cancer risk is associated with lifestyle factors such as diet, smoking, and physical activity, as well as genetic susceptibility.[7] Pancreatic cancer, known for its poor prognosis, is influenced by smoking, chronic pancreatitis, and diabetes.[8] Gallbladder cancer is associated with gallstones, chronic inflammation, and certain infections.[9] The multifaceted nature of these cancers necessitates comprehensive studies to better understand their risk factors and potential preventive measures.[10] Despite advances in early detection and treatment, the overall prognosis of these cancers remains poor. Identifying modifiable risk factors is essential for developing effective prevention strategies.

Circulating blood-based biomarkers have attracted increasing attention in cancer epidemiology because of their accessibility, routine clinical use, and potential mechanistic relevance. Among them, platelets – traditionally known for their role in hemostasis – have also been shown to actively promote tumor progression by releasing cytokines and growth factors that facilitate angiogenesis, immune evasion, and metastasis, particularly in GI cancers.[11] These findings highlight the biological importance of systemic blood components in cancer development. In addition to cellular elements such as platelets, metabolic blood biomarkers such as low-density lipoprotein (LDL) cholesterol and total cholesterol have been linked to membrane structure, cell signaling, and modulation of the tumor microenvironment (TME).[12–14] Serum creatinine and total protein reflect renal function, muscle metabolism, and nutritional status – factors known to influence cancer progression and prognosis.[15–17] Uric acid, while traditionally associated with gout and metabolic disorders, has demonstrated both antioxidant and pro-inflammatory effects that may affect cancer risk in a tissue-specific manner.[18,19] However, evidence from conventional observational studies remains inconclusive because of confounding and potential reverse causation.

In parallel, the causal role of these blood biomarkers in GI cancers has not been comprehensively investigated. Previous studies have often focused on individual biomarkers or single cancer types, limiting the generalizability and comparability of the results. The 4 cancers selected in this study – gastric, colorectal, pancreatic, and gallbladder cancer – are among the most common and lethal GI malignancies worldwide. They also represent distinct anatomical and biological subtypes, providing an opportunity to assess both shared and divergent metabolic risk factors. Furthermore, high-quality genome-wide association study (GWAS) data are available for these cancers, enabling robust Mendelian randomization (MR) analyses.

Mendelian randomization is a genetic epidemiological approach that uses genetic variants as instrumental variables (IVs) to estimate the causal effect of an exposure (e.g., biomarker levels) on disease outcomes.[20,21] Previous MR studies have provided valuable insights into the relationships between various risk factors and diseases, thereby informing clinical practice and public health strategies. For instance, MR has been used to investigate the causal effects of lipid levels, inflammatory markers, and other metabolic factors on cardiovascular disease (CVD), diabetes, and different cancer types.[22–24]

Objectives

The present study aims to evaluate the potential causal effects of 5 clinically relevant blood metabolic biomarkers – LDL cholesterol, total cholesterol, serum creatinine, serum total protein, and uric acid – on the risk of 4 major GI cancers using a 2-sample MR approach. By integrating genetic and epidemiological data, this study seeks to provide novel insights into the metabolic determinants of GI tumorigenesis and to inform future prevention and risk stratification strategies.

Materials and methods

Data source and study design

This study adopted a 2-sample MR design to assess the causal effects of metabolic blood biomarkers on GI cancer risk. Summary-level genetic association data for exposures (LDL cholesterol, total cholesterol, serum creatinine, serum total protein, and uric acid) and outcomes (gastric, colorectal, pancreatic, and gallbladder cancers) were obtained from large-scale GWASs available through the MRC IEU Open GWAS database (https://gwas.mrcieu.ac.uk).

All data were derived from previous studies that had received ethical approval and informed consent.

Selection of genetic instruments and quality control

For each biomarker, single nucleotide polymorphisms (SNPs) strongly associated with the exposure were selected as IVs at the genome-wide significance level ($p < 5 \times 10^{-8}$). To ensure instrument independence, we applied linkage disequilibrium (LD) clumping at $r^2 < 0.001$ using a 10,000-kb window. When fewer SNPs met the strict threshold, a relaxed p-value threshold of 5×10^{-6} was used while retaining LD filtering. Palindromic SNPs with intermediate allele frequencies were excluded. To ensure instrument strength, F-statistics were calculated for each SNP, and variants with $F < 10$ were excluded to avoid weak instrument bias. Additionally, we reported the median and range of F-statistics for each biomarker to characterize the distributional strength of the instruments. These metrics are provided in Supplementary Table 1.

Outcome GWAS inclusion criteria

Summary-level genetic association data for both exposures and outcomes were obtained from the MRC IEU Open GWAS platform. All datasets included in this study were derived exclusively from individuals of European ancestry to minimize population stratification bias. The outcome datasets included large-scale GWASs for gastric cancer (n = 476,116), CRC (n = 470,002), pancreatic cancer (n = 476,245), and gallbladder cancer (n = 907), each comprising more than 24 million SNPs, except for gallbladder cancer, which included approx. 419,000 SNPs.

The selected exposures – LDL cholesterol, total cholesterol, serum total protein, serum creatinine, and serum uric acid – were chosen based on biological plausibility and prior literature and were each associated with more than 19 million SNPs in cohorts exceeding 340,000 individuals. All GWAS datasets underwent rigorous quality control procedures, including imputation to the Haplotype Reference Consortium (HRC) panel and adjustment for population structure. Detailed dataset identifiers, sample sizes, and SNP coverage are provided in Table 1.

Statistical analyses and pleiotropy testing

The estimation of causal effects between blood cell traits and GI tumors was performed using multiple MR methodologies, including inverse variance weighted (IVW), MR-Egger, weighted median, and simple mode methods. The MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) test was employed to detect and adjust for horizontal pleiotropy by identifying and excluding statistical outliers from the analysis. Heterogeneity among the selected genetic instruments was evaluated using Cochran's Q test,

Table 1. Summary of genome-wide association study (GWAS) datasets for exposure and outcome variables

GWAS ID	Exposure	Sample size	Number of SNPs	Population
ebi-a-GCST90018849	gastric cancer	476,116	24,188,662	European
ebi-a-GCST90018808	colorectal cancer	470,002	24,182,361	European
ebi-a-GCST90018893	pancreatic cancer	476,245	24,195,229	European
ebi-a-GCST001404	gallbladder cancer	907	419,385	European
ebi-a-GCST90018961	LDL cholesterol	343,621	19,037,976	European
ebi-a-GCST90018956	HDL cholesterol	315,133	19,051,633	European
ebi-a-GCST90018974	total cholesterol	344,278	19,043,498	European
ebi-a-GCST90025991	serum urea	437,580	4,231,983	European
ebi-a-GCST90025992	serum albumin	400,938	4,219,040	European
ebi-a-GCST90025995	serum total protein	400,482	4,218,824	European
ebi-a-GCST90025948	serum phosphate	400,159	4,218,812	European
ebi-a-GCST90018942	serum alkaline phosphatase	344,292	19,052,566	European
ebi-a-GCST90018979	serum creatinine	344,104	19,034,241	European
ebi-a-GCST90018977	serum uric acid	343,836	19,041,286	European
ebi-a-GCST90092992	total triglycerides	115,082	11,590,399	European

HDL – high-density lipoprotein; LDL – low-density lipoprotein; SNP – single-nucleotide polymorphism.

with p < 0.05 indicating the absence of significant heterogeneity.[25] The presence of pleiotropy was assessed using the MR-Egger intercept test, wherein a nonsignificant p-value suggested no substantial pleiotropy among the IVs. The robustness and reliability of the causal estimates were evaluated using the leave-one-out approach, which entails sequential exclusion of each IV to determine its influence on the overall results. All analyses were performed in R v. 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) using the TwoSampleMR and MR-PRESSO packages.

Results

A summary of the key causal estimates is presented in Table 2.

Gastric cancer

Mendelian randomization analysis (Fig. 1) revealed that genetically predicted LDL cholesterol levels were significantly associated with a decreased risk of gastric cancer (IVW odds ratio (OR) = 0.815, 95% confidence interval (95% CI): 0.703–0.946, p < 0.01). Similarly, total cholesterol also exhibited a protective effect (IVW OR = 0.790, 95% CI: 0.693–0.900, p < 0.01). Serum uric acid was inversely associated with gastric cancer risk (OR = 0.822, 95% CI: 0.719–0.941, p < 0.01). These associations were consistent across the MR-Egger and weighted median methods, although the MR-Egger estimates had wider CIs.

Colorectal cancer

Higher serum total protein levels were associated with a reduced risk of CRC (IVW OR = 0.855, 95% CI: 0.777–0.940, p < 0.01). Serum uric acid also showed a protective effect (OR = 0.855, 95% CI: 0.769–0.951, p < 0.01) (Fig. 1). The robustness of these findings was supported by MR-Egger, weighted median, and MR-PRESSO sensitivity analyses.

Pancreatic cancer

Serum uric acid levels were positively associated with pancreatic cancer risk (IVW OR = 1.246, 95% CI: 1.010–1.539, p = 0.040), contrasting with their protective role

Table 2. Summary of the main Mendelian randomization (MR) findings for biomarkers and gastrointestinal (GI) cancers

Exposure	Cancer outcome	IVW OR (95% CI)	p-value	Effect
LDL cholesterol	gastric cancer	0.815 (0.703–0.946)	<0.01	protective
Total cholesterol	gastric cancer	0.790 (0.693–0.900)	<0.01	protective
Serum creatinine	gallbladder cancer	0.037 (0.005–0.287)	<0.01	strong protective effect
Total protein	colorectal cancer	0.855 (0.777–0.940)	<0.01	protective
Uric acid	colorectal cancer	0.855 (0.769–0.951)	<0.01	protective
Uric acid	gastric cancer	0.822 (0.719–0.941)	<0.01	protective
Uric acid	pancreatic cancer	1.246 (1.010–1.539)	0.040	risk factor

HDL – high-density lipoprotein; LDL – low-density lipoprotein; OR – odds ratio; 95% CI – 95% confidence interval; IVW – inverse variance weighted.

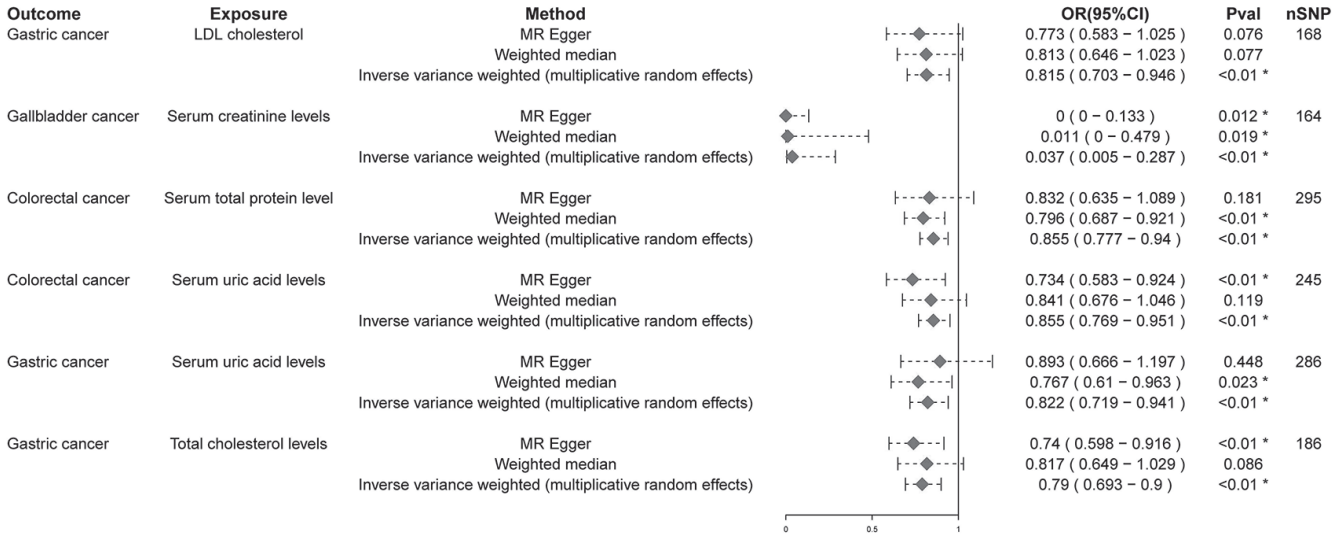


Fig. 1. Association between blood cell traits and GI tumors

GI – gastrointestinal; MR-Egger – Mendelian randomization–Egger; OR – odds ratio; 95% CI – 95% confidence interval; SNP – single nucleotide polymorphism; *indicates statistical significance; LDL – low-density lipoprotein.

in gastric and colorectal cancer (Supplementary Fig. 1). MR-Egger and weighted median estimates demonstrated consistent directionality, although statistical significance was attenuated (weighted median p = 0.061).

Gallbladder cancer

Serum creatinine was strongly associated with a decreased risk of gallbladder cancer (IVW OR = 0.037, 95% CI: 0.005–0.287, p < 0.01), with consistent results across the MR-Egger and weighted median methods (Fig. 1).

Scatter plot visualization

The MR scatter plots presented in Fig. 2 visually support the causal estimates derived from statistical modeling. In Fig. 2A,B, clear negative slopes are observed between LDL cholesterol and gastric cancer, and between serum creatinine and gallbladder cancer, respectively, indicating protective effects. Figure 2C,D shows similar inverse relationships for serum total protein and uric acid with CRC risk. In contrast, Fig. 2F shows a positive slope between uric acid and pancreatic cancer, consistent with its observed risk-enhancing effect. The alignment of most SNPs along the IVW regression lines, with minimal outliers, suggests that the instrumental variables exert consistent directional effects with limited horizontal pleiotropy. These visual findings reinforce the robustness of the causal estimates obtained through MR methods.

Sensitivity and robustness

These estimates were consistent across the MR-Egger and weighted median methods, with no indication of directional pleiotropy (MR-Egger intercept p > 0.05 for all pairs). Heterogeneity testing using Cochran's Q test showed no significant heterogeneity in most models (Q p-value > 0.05) (Table 3). MR-PRESSO analysis did not identify influential outlier SNPs, and leave-one-out analyses further supported the robustness of the results.

The F-statistics of the included instrumental variables were all above the conventional threshold (F > 10), confirming strong instrument strength. For example, in the HDL cholesterol–CRC model, 287 independent SNPs yielded F-statistics with a median of 33.86 (interquartile range (IQR): 26.72–61.65) and a range from 19.74 to 701.81 (Supplementary Table 1).

Discussion

Gastrointestinal cancers, including gastric, colorectal, pancreatic, and gallbladder cancers, are among the leading causes of cancer-related mortality worldwide.[1] The etiology of these cancers is multifactorial, involving genetic predisposition, lifestyle factors, and environmental exposures.[4] Despite advancements in cancer treatment, the prognosis of many GI cancers remains poor, underscoring the importance of identifying modifiable risk factors.[26] This study employed MR to explore the causal relationships between blood cell traits and GI cancer risk, providing insights that could inform clinical practice and public health strategies.

The results of this study show that the protective effect of LDL cholesterol against gastric cancer, as indicated by the IVW method, may be attributed to its role in cellular processes. Low-density lipoprotein cholesterol is known to influence membrane fluidity and signaling pathways that may affect cell proliferation and apoptosis.[12] Furthermore, it is involved in modulation of immune responses, which may contribute to tumor suppression.[13] This finding is consistent with previous studies suggesting that certain lipid profiles may exert protective effects against specific cancers.[14,27]

Serum creatinine levels showed a protective effect against gallbladder cancer. Creatinine, a breakdown product of creatine phosphate derived from muscle metabolism, has been linked to kidney function and overall metabolic health.[15] Its protective role may be related to better metabolic control and lower levels of metabolic waste, thereby reducing chronic inflammation that could contribute to cancer development.[28] This interpretation is supported by studies indicating that improved metabolic health may reduce cancer risk.[29] Given the relatively small sample size of the gallbladder cancer cohort (n = 907), the statistical power to detect modest associations may be limited. Therefore, although the observed protective effect of serum creatinine is biologically plausible, further validation in larger datasets is warranted.

The association between serum total protein levels and a reduced risk of CRC may reflect the overall nutritional status and immune competence of individuals. Higher protein levels may enhance the body's ability to repair DNA damage and maintain immune surveillance, which are crucial in preventing carcinogenesis.[16] Similar mechanisms have been proposed in other studies linking nutritional status to cancer risk.[17]

Serum uric acid levels showed a dual role in cancer risk, acting as a protective factor in colorectal and gastric cancer, but as a risk factor in pancreatic cancer. This duality may stem from uric acid's complex involvement in redox balance and inflammation. At moderate levels, uric acid may act as an antioxidant, whereas at elevated concentrations, it can promote oxidative stress and *NLRP3* inflammasome activation – particularly in pancreatic tissues.[18,19,30,31] These microenvironmental mechanisms may help explain its divergent effects across different cancer types. Furthermore, the association between uric acid and pancreatic cancer might be partially mediated by obesity-related inflammatory pathways,[32] which warrants further evaluation through multivariable MR (MVMR).

Total cholesterol levels showed a significant protective effect against gastric cancer. Cholesterol plays a role in cellular membrane integrity and signaling pathways, and its

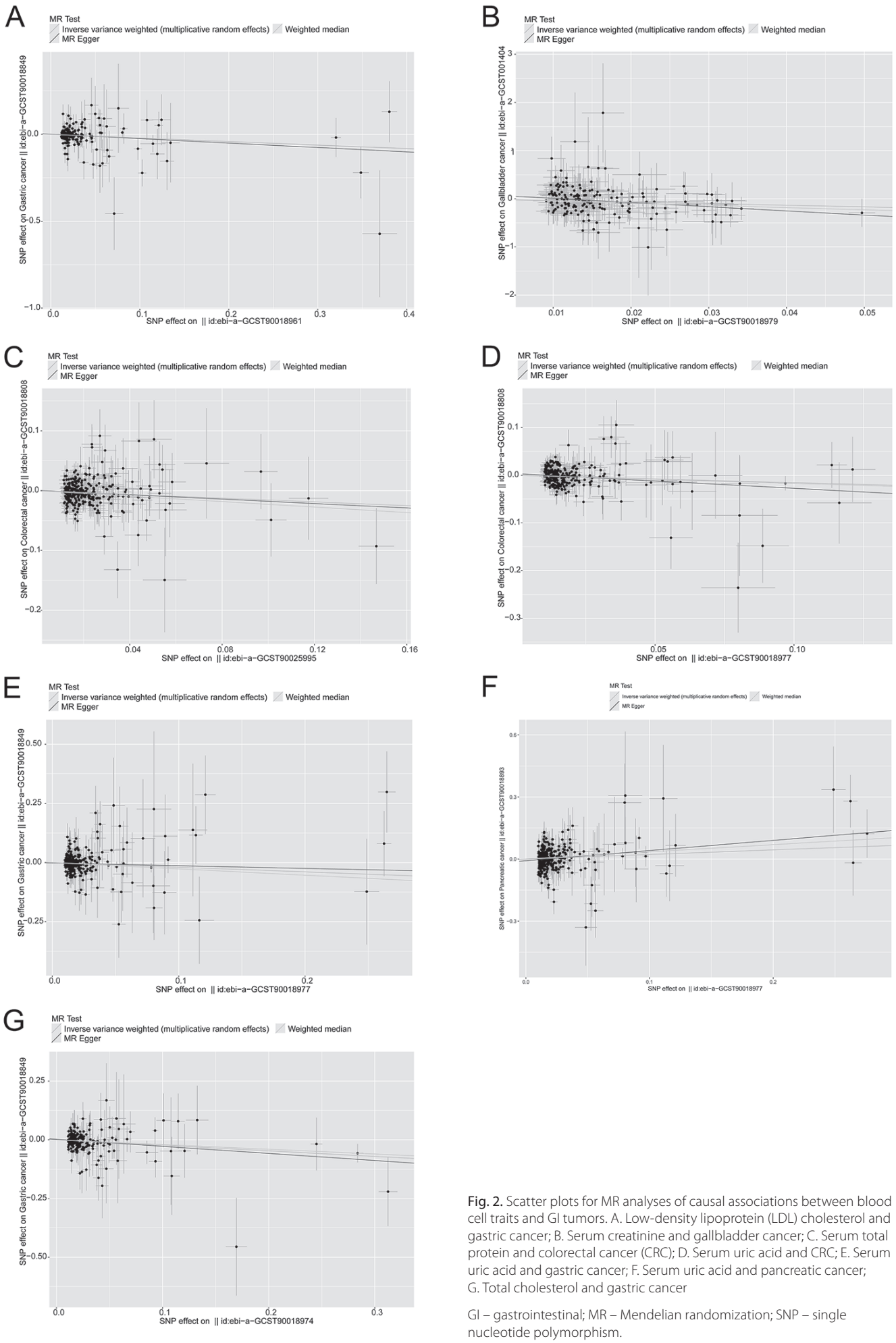


Fig. 2. Scatter plots for MR analyses of causal associations between blood cell traits and GI tumors. A. Low-density lipoprotein (LDL) cholesterol and gastric cancer; B. Serum creatinine and gallbladder cancer; C. Serum total protein and colorectal cancer (CRC); D. Serum uric acid and CRC; E. Serum uric acid and gastric cancer; F. Serum uric acid and pancreatic cancer; G. Total cholesterol and gastric cancer

GI – gastrointestinal; MR – Mendelian randomization; SNP – single nucleotide polymorphism.

Table 3. Assessment of pleiotropy and heterogeneity in Mendelian randomization (MR) models across biomarkers and cancer outcomes

Outcome	Exposure	Heterogeneity		Pleiotropy
		Q	Q p-value	p-value
Colorectal cancer	HDL cholesterol	316.383	0.097	0.197
Gallbladder cancer	HDL cholesterol	135.428	0.144	0.464
Pancreatic cancer	HDL cholesterol	328.424	0.171	0.960
Colorectal cancer	LDL cholesterol	150.300	0.301	0.744
Gallbladder cancer	LDL cholesterol	68.786	0.157	0.948
Gastric cancer	LDL cholesterol	185.351	0.145	0.662
Pancreatic cancer	LDL cholesterol	180.006	0.322	0.443
Colorectal cancer	serum albumin	269.259	0.087	0.087
Gallbladder cancer	serum albumin	98.235	0.918	0.624
Gastric cancer	serum albumin	244.249	0.359	0.921
Pancreatic cancer	serum albumin	278.128	0.107	0.081
Colorectal cancer	serum alkaline phosphatase	326.321	0.333	0.505
Gallbladder cancer	serum alkaline phosphatase	94.128	0.961	0.666
Gastric cancer	serum alkaline phosphatase	394.954	0.093	0.310
Pancreatic cancer	serum alkaline phosphatase	311.963	0.675	0.371
Colorectal cancer	serum creatinine	422.894	0.197	0.404
Gallbladder cancer	serum creatinine	102.047	1.000	0.089
Gastric cancer	serum creatinine	455.188	0.150	0.646
Pancreatic cancer	serum creatinine	436.112	0.583	0.624
Colorectal cancer	serum phosphate	200.221	0.169	0.513
Gallbladder cancer	serum phosphate	108.639	0.088	0.950
Gastric cancer	serum phosphate	190.015	0.327	0.919
Pancreatic cancer	serum phosphate	215.832	0.125	0.407
Colorectal cancer	serum total protein	290.537	0.530	0.830
Gallbladder cancer	serum total protein	177.126	0.080	0.872
Pancreatic cancer	serum total protein	330.559	0.497	0.381
Colorectal cancer	serum urea	218.545	0.076	0.217
Gallbladder cancer	serum urea	83.991	0.804	0.619
Gastric cancer	serum urea	231.534	0.083	0.372
Pancreatic cancer	serum urea	228.816	0.112	0.541
Colorectal cancer	serum uric acid	223.663	0.808	0.140
Gallbladder cancer	serum uric acid	109.855	0.618	0.723
Gastric cancer	serum uric acid	286.208	0.452	0.538
Pancreatic cancer	serum uric acid	283.044	0.666	0.054
Colorectal cancer	total cholesterol	197.675	0.096	0.582
Gallbladder cancer	total cholesterol	80.917	0.246	0.444
Gastric cancer	total cholesterol	202.338	0.168	0.451
Pancreatic cancer	total cholesterol	229.433	0.090	0.736
Colorectal cancer	total triglycerides	82.093	0.075	0.066
Gallbladder cancer	total triglycerides	8.269	0.974	0.079
Gastric cancer	total triglycerides	72.943	0.208	0.390
Pancreatic cancer	total triglycerides	68.491	0.529	0.253

HDL – high-density lipoprotein; LDL – low-density lipoprotein.

modulation may affect cancer cell behavior.[33] The protective effect observed in this study aligns with previous findings suggesting that maintaining balanced cholesterol levels may contribute to cancer prevention.[34]

The observed associations between blood cell traits and GI cancers have significant implications for clinical practice. For instance, monitoring and potentially modifying lipid profiles, creatinine, total protein, and uric acid levels in patients could represent a strategy to mitigate cancer risk.[35] The findings suggest that clinicians should consider the impact of these blood cell traits on cancer risk, especially in patients with a high predisposition to GI cancers.[36] Among these associations, the strong inverse association observed between serum creatinine and gallbladder cancer highlights potential biological relevance. However, its clinical utility is limited by the low incidence of the disease and the lack of established screening strategies. Further studies are needed to confirm this association and explore its feasibility for individualized risk prediction.

Limitations of the study

Several limitations should be acknowledged. First, the scope of blood biomarkers analyzed was limited. Other potentially relevant variables such as inflammatory markers (e.g., C-reactive protein (CRP)), micronutrients (e.g., vitamin D), or liver enzymes (e.g., alanine transaminase (ALT) and aspartate transaminase (AST)) were not included because of data availability or the lack of strong genetic instruments. Second, we did not perform MVMR, which may help disentangle collinearity between closely related exposures such as LDL and total cholesterol. This should be prioritized in future studies. Third, our GWAS data were derived exclusively from individuals of European ancestry, which may limit generalizability to other ethnic groups. This is particularly relevant for cancers such as gastric and gallbladder cancer, which have higher incidence rates in Asian and South American populations. Trans-ethnic validation will be essential to establish broader applicability.

Additionally, we did not examine sex- or age-specific effects, nor did we stratify by cancer subtype (e.g., colon compared with rectal cancer), which may obscure biologically meaningful heterogeneity. Future work could incorporate stratified MR analyses to explore population-specific and tumor-specific metabolic pathways. Finally, functional validation using experimental models is needed to confirm the mechanistic pathways suggested by our findings.

While our findings suggest that certain metabolic blood biomarkers, such as LDL cholesterol and serum uric acid, may play protective roles in specific GI cancers, these biomarkers are also known risk factors for CVD, gout, and kidney dysfunction. Thus, it is essential to interpret our findings within a broader clinical context. We do not advocate artificial elevation of these biomarkers as a preventive strategy. Instead, the observed associations may reflect complex metabolic or immunological mechanisms that warrant further mechanistic investigation. Clinical interventions should always consider the balance between cancer risk and other disease burdens.

Conclusions

This study provides genetic evidence supporting causal links between selected blood metabolic traits and GI cancer risk. These results highlight new avenues for biomarker-guided screening and underscore the need for further mechanistic and translational research to translate these associations into actionable clinical tools.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15621795. The package contains the following files:

Supplementary Table 1. Instrumental variables for biomarkers in GI cancers.

Supplementary Fig. 1. The detailed results from the MR analysis for all identified markers and their associations with GI cancers.

Data Availability Statement

Data sharing does not apply to this article, as no new data were generated during the current study.

Consent of publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Xiufeng Huang https://orcid.org/0009-0009-7979-925X
Jianxin Chen https://orcid.org/0000-0003-2467-3333
Feiteng Gu https://orcid.org/0000-0002-2789-074X
Xiaodan Zheng https://orcid.org/0009-0003-2731-1888

References

1. Dizdar Ö, Kılıçkap S. Global epidemiology of gastrointestinal cancers. In: Yalçin Ş, Philip PA, eds. *Textbook of Gastrointestinal Oncology*. Cham, Switzerland: Springer; 2019:1–12. ISBN:978-3-030-18888-7.
2. Wang S, Zheng R, Li J, et al. Global, regional, and national lifetime risks of developing and dying from gastrointestinal cancers in 185 countries: A population-based systematic analysis of GLOBOCAN. *Lancet Gastroenterol Hepatol*. 2024;9(3):229–237. doi:10.1016/S2468-1253(23)00366-7
3. Huang M, Feng L, Ren H, et al. Knowledge, attitudes, and practices regarding whole-course management among patients with gastrointestinal cancers: A cross-sectional study. *World J Surg Oncol*. 2025;23(1):45. doi:10.1186/s12957-025-03668-7

4. Ashktorab H, Kupfer SS, Brim H, Carethers JM. Racial disparity in gastrointestinal cancer risk. *Gastroenterology*. 2017;153(4):910–923. doi:10.1053/j.gastro.2017.08.018
5. Barone E, Corrado A, Gemignani F, Landi S. Environmental risk factors for pancreatic cancer: An update. *Arch Toxicol*. 2016;90(11): 2617–2642. doi:10.1007/s00204-016-1821-9
6. Yang L, Ying X, Liu S, et al. Gastric cancer: Epidemiology, risk factors and prevention strategies. *Chin J Cancer Res*. 2020;32(6):695–704. doi:10.21147/j.issn.1000-9604.2020.06.03
7. Lewandowska A, Rudzki G, Lewandowski T, Stryjkowska-Góra A, Rudzki S. Risk factors for the diagnosis of colorectal cancer. *Cancer Control*. 2022;29:10732748211056692. doi:10.1177/10732748211056692
8. Hu JX, Zhao CF, Chen WB, et al. Pancreatic cancer: A review of epidemiology, trend, and risk factors. *World J Gastroenterol*. 2021;27(27): 4298–4321. doi:10.3748/wjg.v27.i27.4298
9. Pérez-Moreno P, Riquelme I, García P, Brebi P, Roa JC. Environmental and lifestyle risk factors in the carcinogenesis of gallbladder cancer. *J Pers Med*. 2022;12(2):234. doi:10.3390/jpm12020234
10. Drewes AM, Olesen AE, Farmer AD, Szigethy E, Rebours V, Olesen SS. Gastrointestinal pain. *Nat Rev Dis Primers*. 2020;6(1):1. doi:10.1038/s41572-019-0135-7
11. Bianchi S, Torge D, Rinaldi F, Piattelli M, Bernardi S, Varvara G. Platelets' role in dentistry: From oral pathology to regenerative potential. *Biomedicines*. 2022;10(2):218. doi:10.3390/biomedicines10020218
12. Codini M, Garcia-Gil M, Albi E. Cholesterol and sphingolipid enriched lipid rafts as therapeutic targets in cancer. *Int J Mol Sci*. 2021;22(2):726. doi:10.3390/ijms22020726
13. Huang B, Song BL, Xu C. Cholesterol metabolism in cancer: Mechanisms and therapeutic opportunities. *Nat Metab*. 2020;2(2):132–141. doi:10.1038/s42255-020-0174-0
14. Saini RK, Rengasamy KRR, Mahomoodally FM, Keum YS. Protective effects of lycopene in cancer, cardiovascular, and neurodegenerative diseases: An update on epidemiological and mechanistic perspectives. *Pharmacol Res*. 2020;155:104730. doi:10.1016/j.phrs.2020.104730
15. De Rosa S, Greco M, Rauseo M, Annetta MG. The good, the bad, and the serum creatinine: Exploring the effect of muscle mass and nutrition. *Blood Purif*. 2023;52(9–10):775–785. doi:10.1159/000533173
16. Molinaro C, Martoriati A, Cailliau K. Proteins from the DNA damage response: Regulation, dysfunction, and anticancer strategies. *Cancers (Basel)*. 2021;13(15):3819. doi:10.3390/cancers13153819
17. Bose S, Allen AE, Locasale JW. The molecular link from diet to cancer cell metabolism. *Mol Cell*. 2020;80(3):554. doi:10.1016/j.molcel.2020.10.006
18. Gherghina ME, Peride I, Tiglis M, Neagu TP, Niculae A, Checherita IA. Uric acid and oxidative stress: Relationship with cardiovascular, metabolic, and renal impairment. *Int J Mol Sci*. 2022;23(6):3188. doi:10.3390/ijms23063188
19. Mi S, Gong L, Sui Z. Friend or foe? An unrecognized role of uric acid in cancer development and the potential anticancer effects of uric acid-lowering drugs. *J Cancer*. 2020;11(17):5236–5244. doi:10.7150/jca.46200
20. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Methods Primers*. 2022;2(1):6. doi:10.1038/s43586-021-00092-5
21. Richmond RC, Davey Smith G. Mendelian randomization: Concepts and scope. *Cold Spring Harb Perspect Med*. 2022;12(1):a040501. doi:10.1101/cshperspect.a040501
22. Yuan S, Carter P, Bruzelius M, et al. Effects of tumour necrosis factor on cardiovascular disease and cancer: A two-sample Mendelian randomization study. *eBioMedicine*. 2020;59:102956. doi:10.1016/j.ebiom.2020.102956
23. Chen K, Li J, Ouyang Y, et al. Blood lipid metabolic profiles and causal links to site-specific cancer risks: A Mendelian randomization study. *Nutr Cancer*. 2024;76(2):175–186. doi:10.1080/01635581.2023.2294521
24. Wang Z, Chen J, Zhu L, Jiao S, Chen Y, Sun Y. Metabolic disorders and risk of cardiovascular diseases: A two-sample mendelian randomization study. *BMC Cardiovasc Disord*. 2023;23(1):529. doi:10.1186/s12872-023-03567-3
25. Levin MG, Burgess S. Mendelian randomization as a tool for cardiovascular research: A review. *JAMA Cardiol*. 2024;9(1):79. doi:10.1001/jamacardio.2023.4115
26. Jardim SR, De Souza LMP, De Souza HSP. The rise of gastrointestinal cancers as a global phenomenon: Unhealthy behavior or progress? *Int J Environ Res Public Health*. 2023;20(4):3640. doi:10.3390/ijerph20043640
27. Martin-Perez M, Urdiroz-Urricelqui U, Bigas C, Benitah SA. The role of lipids in cancer progression and metastasis. *Cell Metab*. 2022;34(11): 1675–1699. doi:10.1016/j.cmet.2022.09.023
28. Peris-Fernández M, Roca-Marugán M, Amengual JL, et al. Uremic toxins and inflammation: Metabolic pathways affected in non-dialysis-dependent stage 5 chronic kidney disease. *Biomedicines*. 2024;12(3):607. doi:10.3390/biomedicines12030607
29. Clemente-Suárez VJ, Martín-Rodríguez A, Redondo-Flórez L, et al. Metabolic health, mitochondrial fitness, physical activity, and cancer. *Cancers (Basel)*. 2023;15(3):814. doi:10.3390/cancers15030814
30. Liu Y, Liu Y, Wang Q, et al. MIF inhibitor ISO-1 alleviates severe acute pancreatitis-associated acute kidney injury by suppressing the NLRP3 inflammasome signaling pathway. *Int Immunopharmacol*. 2021;96:107555. doi:10.1016/j.intimp.2021.107555
31. Braga TT, Forni MF, Correa-Costa M, et al. Soluble uric acid activates the NLRP3 inflammasome. *Sci Rep*. 2017;7(1):39884. doi:10.1038/srep39884
32. Park SK, Oh C, Kim M, Ha E, Choi Y, Ryoo J. Metabolic syndrome, metabolic components, and their relation to the risk of pancreatic cancer. *Cancer*. 2020;126(9):1979–1986. doi:10.1002/cncr.32737
33. Maja M, Tyteca D. Alteration of cholesterol distribution at the plasma membrane of cancer cells: From evidence to pathophysiological implication and promising therapy strategy. *Front Physiol*. 2022;13:999883. doi:10.3389/fphys.2022.999883
34. Xu H, Zhou S, Tang Q, Xia H, Bi F. Cholesterol metabolism: New functions and therapeutic approaches in cancer. *Biochim Biophys Acta Rev Cancer*. 2020;1874(1):188394. doi:10.1016/j.bbcan.2020.188394
35. Dent SF, Kikuchi R, Kondapalli L, et al. Optimizing cardiovascular health in patients with cancer: A practical review of risk assessment, monitoring, and prevention of cancer treatment-related cardiovascular toxicity. *Am Soc Clin Oncol Educ Book*. 2020;(40):501–515. doi:10.1200/EDBK_286019
36. Zametkin E, Guyer D, Tarshish Y, Bash K, Almhanna K. Total parenteral nutrition for patients with gastrointestinal cancers: A clinical practice review. *Ann Palliat Med*. 2023;12(5):1072–1080. doi:10.21037/apm-22-1380

miR-145 inhibits colorectal cancer progression and metastasis by targeting *SNAI1* and the cAMP/PKA pathway

Jianshan Cai[1,A,D], Qiang Sun[2,C], Shichao Deng[2,B,D], Qi Wei[2,B], Longzhi Li[2,B], Baojin Ma[2,B], Jiadong Chen[2,A,F]

[1] Department of General Surgery, Huashan Hospital, Fudan University, Shanghai, China
[2] Department of General Surgery, Jing'an District Center Hospital of Shanghai, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1047–1061

Address for correspondence
Jiadong Chen
E-mail: chenjiadong8@sina.com

Funding sources
None declared

Conflict of interest
None declared

Received on April 11, 2025
Reviewed on May 9, 2025
Accepted on September 12, 2025

Published online on June 29, 2026

Abstract

Background. Transcriptional repressor 1 of the Snail family (*SNAI1*) and miR-145 play crucial roles in the development of colon cancer. Tumor invasiveness and epithelial–mesenchymal transition (EMT) are associated with *SNAI1*, whereas miR-145 is considered a tumor suppressor. Although cAMP/PKA signaling is critical for regulating cell survival and invasiveness, the precise mechanisms through which *SNAI1* and miR-145 influence tumor cell behavior via this pathway remain unclear.

Objectives. To examine the mechanisms by which *SNAI1* and miR-145 regulate colon cancer cell invasion and survival through the cAMP/PKA signaling pathway and to assess their functional relationship.

Materials and methods. *SNAI1* and miR-145 were overexpressed in the colon cancer cell lines DLD1 and HCT116, respectively, and cAMP and PKA activity was assessed using quantitative real-time PCR (qPCR) and western blot analysis. In addition, a luciferase reporter assay and application of the PKA inhibitor H89 were used to further investigate the roles of *SNAI1* and miR-145 in regulating cell invasion and survival.

Results. Overexpression of *SNAI1* significantly increased cAMP and PKA activity and promoted cell survival and invasion. However, overexpression of miR-145 decreased cell invasion and survival by blocking the cAMP/PKA signaling pathway. When both were overexpressed, miR-145 further decreased cAMP/PKA activity and inhibited the tumor-promoting effects of *SNAI1*. Use of the PKA inhibitor H89 also reduced the tumor-promoting effect of *SNAI1*.

Conclusions. By activating the cAMP/PKA signaling pathway, *SNAI1* promotes colon cancer cell survival and invasion, whereas miR-145 suppresses tumor progression by inhibiting this pathway. These findings suggest that *SNAI1* and miR-145 may represent viable therapeutic targets for colon cancer treatment and that targeting the cAMP/PKA pathway may provide a novel therapeutic strategy.

Key words: colorectal neoplasms, microRNA-145, Snail family transcription factors, cyclic AMP-dependent protein kinases, signal transduction

Cite as
Cai J, Sun Q, Deng S, et al. miR-145 inhibits colorectal cancer progression and metastasis by targeting *SNAI1* and the cAMP/PKA pathway. *Adv Clin Exp Med*. 2026;35(6):1047–1061. doi:10.17219/acem/210638

DOI
10.17219/acem/210638

Highlights

- SNAI1 promotes colon cancer cell survival and invasion by activating the cAMP/PKA signaling pathway, a key regulator of tumor progression.
- Tumor suppressor miR-145 inhibits colon cancer invasiveness and cell survival through suppression of cAMP/PKA signaling.
- miR-145 counteracts the oncogenic effects of SNAI1, reducing cAMP/PKA activity and limiting cancer cell aggressiveness.
- Targeting the SNAI1–miR-145–cAMP/PKA regulatory axis may offer a novel therapeutic strategy for colorectal cancer treatment.

Background

With the acceleration of globalization, population aging and urbanization, global public, health challenges have become more complex. Among these challenges, the high incidence and mortality rates of malignant tumors are particularly concerning. Colorectal cancer (CRC) is a typical representative among them. It ranks 3rd in terms of new cases and 2nd in terms of mortality among all malignant tumors, with around 1.88 million new cases and 920,000 deaths, according to the 2020 global cancer statistics.[1] Although advancements in screening and treatment technologies have led to a decline in CRC incidence and mortality in several high-income countries, the rates continue to rise in many developing nations, particularly in China.[1]

About 60% of patients receive an advanced diagnosis of CRC because of the disease's complicated molecular features and the absence of clear early signs.[2] Colorectal cancer is characterized by significant molecular heterogeneity, which contributes to its varied clinical presentation and delayed diagnosis. This complexity is driven by a high frequency of genetic mutations, such as alterations in *APC*, *KRAS*, *TP53*, and *PIK3CA*, that promote tumor initiation and progression. In addition to genetic changes, epigenetic modifications – such as DNA methylation abnormalities and histone modifications – further disrupt the regulation of gene expression. Moreover, multiple signaling pathways, including Wnt/β-catenin, MAPK, PI3K/AKT, and tumor growth factor beta (TGF-β), are frequently dysregulated, enhancing tumor cell proliferation, invasion, and metastasis. The interplay of these molecular alterations results in substantial tumor heterogeneity, complicating early detection and the development of standardized treatment strategies. Together with the absence of apparent early symptoms, these molecular features contribute to the high proportion of CRC patients diagnosed at advanced stages.

Currently, the primary treatment options for CRC treatments remain conventional approaches such as radiotherapy, chemotherapy, and surgical resection. However, the effect of these treatments is restricted by many factors, especially the existence of cancer stem cells (CSCs), which seriously affects the therapeutic effect. These cells are a small fraction of cells within a tumor with high self-renewal ability, which can re-establish the tumor after treatment, and may acquire stronger malignant properties after radiotherapy and chemotherapy, leading to cancer recurrence and metastasis.[2,3]

In addition, the epithelial–mesenchymal transition (EMT) is a significant contributing factor to the development of CRC. The process of EMT gives tumor cells the capacity to invade and migrate, which significantly increases the likelihood that the tumor will spread.[4] In this process, transcription repressor 1 of Snail family (*SNAI1*, or Snail) is an important regulatory factor.[5] It enhances tumor cells' capacity to migrate, which in turn facilitates their invasion, by suppressing the expression of adhesion molecules on the surface of epithelial cells.[5]

MicroRNAs (miRNAs) are a class of small noncoding RNAs that play a key role in regulating the expression of genes associated with malignancies at the molecular level.[6] Among these, miR-145 has been shown to suppress tumor growth in a variety of cancers.[7] By downregulating the expression of genes associated with tumor proliferation and metastasis, such as *IGF1R*, fascin-1, and paxillin, miR-145 can inhibit CRC growth and metastasis.[7,8] Furthermore, miR-145 regulates the p53 tumor suppressor gene network, which may significantly increase tumor responsiveness to treatment and improve therapeutic outcomes.[9]

Although previous studies have provided some understanding of the roles of CSCs, EMT, and miR-145 in CRC, the mechanisms through which these molecules and cellular processes interact, and how they jointly drive tumor progression and treatment resistance, remain incompletely understood.

Objectives

The aim of this study was to explore the regulatory roles of *SNAI1* and miR-145 in CRC cell survival and invasion, as well as their potential significance for therapeutic strategies, in order to identify new targets and strategies for CRC treatment.

Materials and methods

Mining of data

To explore the expression levels of EMT transcription factors in human rectal cancer, the colorectal adenocarcinoma dataset from The Cancer Genome Atlas (TCGA) was analyzed using cBioPortal for Cancer Genomics (https://www.cbioportal.org) to determine the frequency of *SNAI1*, *SNAI2*, *ZEB1*, and *ZEB2* gene alterations. Simultaneously, the public dataset available in Research Edition Oncomine 4.4 (Compendia Bioscience, Ann Arbor, USA; https://www.oncomine.org) was used to extract the expression values of *SNAI1*, *SNAI2*, *ZEB1*, and *ZEB2* specifically in rectal cancer. The following filtering criteria were applied to identify the relevant datasets: genes: *ZEB1*, *ZEB2*, *SNAI1*, and *SNAI2*; analysis type: cancer vs normal; cancer type: rectal cancer. The datasets selected for analysis and visualization were the largest complete datasets available and were ranked according to overexpression p-values.

Cell lines

Human CRC cell lines (e.g., HCT116 and DLD1) were obtained from the American Type Culture Collection (ATCC; Manassas, USA). Cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Gibco, Waltham, USA) supplemented with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin–streptomycin (Gibco). Cultures were maintained at 37°C in a humidified atmosphere containing 5% CO_2. All procedures were performed in accordance with ATCC guidelines.

Overexpression of *SNAI1*

As directed by the manufacturer, stable cell lines overexpressing *SNAI1* were generated using Lipofectamine 3000 (Invitrogen, Carlsbad, USA). Empty vectors (pCMV-3Tag-1) and pCMV-*SNAI1* were transfected into HCT116 and DLD1 cells to generate the stable cell lines HCT116-*SNAI1* and DLD1-*SNAI1*, respectively. Transfected cells were selected using medium containing G418 at a final concentration of 400 mg/mL (Invitrogen). *SNAI1* expression was verified with western blot analysis.

Overexpression of miR-145

HCT116 and DLD1 cell lines were separately seeded in 6-well plates and cultured to 70–80% confluence. Colorectal cancer cells were transfected using Lipofectamine 3000 (Invitrogen) transfection reagent in accordance with the manufacturer's instructions, using either miR-145 mimics (synthetic miR-145 RNA molecules) or negative-control mimics (NC mimics). Following transfection, the cells were cultured in fresh medium for an additional 48 h to assess miR-145 expression levels and their impact on the biological properties of the cells. To confirm successful transfection and efficient overexpression of miR-145 mimics, quantitative real-time polymerase chain reaction (qPCR) was used to evaluate miR-145 expression levels in the transfected cells.

Transwell assay

Cell migration was assessed using 24-well Transwell (Boyden) chambers with 8-μm pore-size inserts (Corning Company, Corning, USA), and the assays were performed according to the manufacturer's protocol. Cells adhering to the underside of each insert membrane were stained with crystal violet and subsequently counted using a low-power Leica microscope (M50 stereomicroscope; Leica Camera AG, Wetzlar, Germany).

Colony formation

After seeding into 6-well plates, DLD1-*SNAI1* and HCT116-*SNAI1* cells were compared with empty-vector control cells. The plates received 0, 2, 4, and 6 Gy X-ray radiation, respectively, delivered using a Precision X-ray irradiator (Model X-RAD 320; Precision X-Ray, Inc., North Branford, USA). In a 2nd experiment, DLD1-*SNAI1* and HCT116-*SNAI1* cells were transfected with 2.5 μg of pEZX vector (Genecopoeia, Rockville, USA) containing either a scrambled control (SCR) sequence or a miR-145 expression construct. Lipofectamine 3000 (Invitrogen) was used for transfection.

After irradiation, the cells were replated in triplicate in 6-well plates at densities of 400, 800, 1000, and 1200 cells per well to determine survival scores. Following incubation for 10 days, the cells underwent 3 washes with phosphate-buffered saline (PBS), fixation in 70% ethanol, and staining with 0.5% crystal violet. After staining, the number of colonies in each well was counted under a light microscope (M50; Leica Microsystems) and recorded for subsequent analysis.

Luciferase promoter detection

Yin-Yuan Mo, PhD, from the University of Mississippi (Oxford, USA) kindly donated a luciferase reporter plasmid containing a putative 1.4 kb miR-145 promoter cloned into the pGL3 basic vector (Promega, Madison, USA). Luciferase activity was assessed in stable empty-vector-transfected cells, DLD1-*SNAI1* cells, and HCT116-*SNAI1* cells. The miR-145 promoter construct and the Renilla luciferase plasmid were co-transfected into cells cultured in 12-well plates. Briefly, 48 h after transfection, the cells were lysed for the luciferase assay. The luciferase assay was performed using a dual-luciferase assay kit (Promega) in accordance with the manufacturer's instructions.

Table 1. Primer sequences used for quantitative real-time polymerase chain reaction (qPCR) analysis

Gene	Forward primer (5′ → 3′)	Reverse primer (5′ → 3′)
Snail	TGCCCTCAAGATGCACATCCGA	GGGACAGGAGAAGGGCTTCTC
miR-145	GGTCCAGTTTTCCCAGG	CAGTGCGTGTCGTGGAGT
cAMP	GAACCGCAGTATCATGCTGG	TCCTTGAATTAAGCCGTTCATCA
PKA	GATTCAGACTCGGATTGCTAACG	ACCACTTATGAGCCACTCTACTT
GAPDH	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA

cAMP assay

HCT116 and DLD1 cells were cultured to logarithmic growth phase. The cells were transfected with either miR-145 mimics or NC mimics. To evaluate changes in cAMP levels, miR-145-overexpressing cells were treated with cAMP activators (e.g., forskolin) or cAMP inhibitors (e.g., IBMX). Intracellular cAMP levels were measured using a cAMP enzyme-linked immunosorbent assay (ELISA) kit in accordance with the manufacturer's instructions. The ELISA method was used to quantify cAMP levels based on a colorimetric reaction, and the data were converted to specific concentrations using a standard curve.

Determination of PKA activity

HCT116 and DLD1 cells were transfected with either miR-145 mimics or NC, as described for the cAMP assay. Following a 24-h treatment period, the cells were stimulated with agents such as IBMX or forskolin to induce an increase in cAMP levels. Intracellular PKA activity was measured using a PKA (Protein Kinase A) Colorimetric Activity Kit (cat. No. EIAPKA; Thermo Fisher Scientific, Waltham, USA). The absorbance was measured using a 96-well microplate reader at 450 nm, with 650 nm as the reference wavelength. PKA activity based on the kit standard curve. To further confirm activation of the PKA signaling pathway, CREB phosphorylation levels and other downstream PKA markers were analyzed using western blot analysis.

Quantitative PCR and RNA extraction

Total RNA was extracted from cultured cells using the RNeasy Plus Mini Kit (Qiagen, Valencia, USA). Total RNA (0.8 μg) was reverse transcribed using iScript (Bio-Rad, Hercules, USA) in a 20-μL reaction volume. Quantitative real-time PCR was performed using the Roche LightCycler 480 system and the Universal Probe Library (UPL) monochrome probe system (Roche Diagnostics, Basel, Switzerland), with 5 μL of 1:16 diluted cDNA from each cell-line sample. After 10 min of pre-incubation at 95°C, each gene underwent 55 cycles consisting of 15 s of denaturation at 95°C and 30 s of amplification/extension at 60°C. After completion of the cycling protocol, samples were cooled at 40°C for 30 s. Each cDNA sample was analyzed in triplicate. *GAPDH* was used as the reference gene for data normalization, and multiple samples were analyzed for validation. Primer sequences are shown in Table 1.

Enzyme-linked immunosorbent assay

GRIK5-overexpressing CT26/HCT116 cells or vector control cells (1×10^6) were collected and lysed using cell lysis buffer. The cAMP Detection Kit (Abcam, Cambridge, UK) was then used to measure cAMP levels in the supernatant in accordance with the manufacturer's protocol. Each experiment was performed in at least 3 independent runs.

Western blot analysis

A protease inhibitor cocktail (Sigma-Aldrich, St. Louis, USA) was added to the radioimmunoprecipitation assay (RIPA) lysis buffer in which the cells were suspended (Thermo Fisher Scientific). Protein concentration was determined using the bicinchoninic acid (BCA) protein quantification kit (Thermo Fisher Scientific). Total protein (50 μg) was separated using sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) on a 10% polyacrylamide gel and transferred onto a nitrocellulose membrane (GE Healthcare, Marburg, USA). Primary antibodies against E-cadherin (BD Biosciences, San Diego, USA), *SNAI1*, *SNAI2*, c-Myc, Nanog, Vimentin, KLF4, Sox2, Oct4, and actin (Cell Signaling Technology (CST), San Diego, USA), as well as horseradish peroxidase (HRP)-conjugated secondary antibodies (Jackson ImmunoResearch Laboratories, West Grove, USA), were used for membrane detection. Immunoblots were visualized using enhanced chemiluminescence (GE Healthcare) and the LI-COR Odyssey imaging system (LI-COR BioSciences, Lincoln, USA).

Statistical analyses

GraphPad Prism v. 8.0.2 (GraphPad Software, San Diego, USA) and IBM SPSS v. 25 (IBM Corp., Armonk, USA) were used for statistical analysis of the data in each group. When the sample size was 10–50, the Shapiro–Wilk test was used to assess normality of the correlation data. When the sample size was >50, the Kolmogorov–Smirnov test was used to assess normality of the correlation data, and the relevant results are detailed in Supplementary Table 1.

For data conforming to a normal distribution, means and standard deviations (SDs) are presented in the tables. When the results conformed to a normal distribution, homogeneity of variance was further assessed using the Brown–Forsythe test, and the specific results are listed in Supplementary Table 2. Next, when the sample size was greater than 10 and the results were consistent with normal distribution and homogeneity of variance, one-way analysis of variance (ANOVA) was used to detect overall differences among multiple groups, followed by Tukey's post hoc test to determine which groups differed significantly. When the sample size was less than 10, the Mann–Whitney U test was used for nonparametric comparisons between 2 groups. Nonparametric comparisons among multiple groups were performed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction. The number of corrections corresponded to the total number of pairwise group comparisons, and the specific results are listed in Supplementary Table 3. Statistical significance was defined as $p < 0.05$.

Results

Expression and genetic variation of *SNAI1* and *SNAI2* in patients with colon cancer

Four colon cancer datasets were analyzed to compare the expression levels of *SNAI1* and *SNAI2*. In the Gaedcke cohort (65 patients), *SNAI1* expression was significantly increased and was the highest among the 4 datasets, whereas *SNAI2* expression was significantly decreased, suggesting that *SNAI1* upregulation and *SNAI2* suppression may be associated with aggressive tumor behavior. In the TCGA cohort (60 patients), the results showed decreased *SNAI1* expression and significant inhibition of *SNAI2*, indicating a potential shift in transcriptional regulation during tumor progression. In the Kurashina cohort (37 patients), *SNAI1* expression was the lowest, whereas *SNAI2* expression was higher. In the TCGA 2 cohort (90 patients), the expression patterns of *SNAI1* and *SNAI2* were similar to those observed in the Kurashina cohort (Fig. 1A). These findings imply that *SNAI1* and *SNAI2* expression varies significantly among cohorts.

Further analysis of genetic alterations in *SNAI1* and *SNAI2* in the TCGA dataset showed that the *SNAI1* gene amplification rate was significantly higher than that of *SNAI2* (Fig. 1B), further supporting the dominant role of *SNAI1* in CRC progression. Overexpression of *SNAI1* in HCT116 and DLD1 cell lines was verified using qPCR analysis. In HCT116-*SNAI1* and DLD1-*SNAI1* cells, *SNAI1* mRNA levels were significantly increased, indicating successful overexpression (Fig. 1C). Western blot analysis further confirmed high *SNAI1* overexpression at the protein level, with significantly increased *SNAI1* protein levels in HCT116-*SNAI1* and DLD1-*SNAI1* cells (Fig. 1D), providing a robust experimental basis for subsequent functional studies.

SNAI1's function in colon cancer cell migration, survival, proliferation, and phenotypic transformation

We compared the proliferative capacity of control (Vec) cells and *SNAI1*-overexpressing colon cancer cell lines HCT116 and DLD1. The findings demonstrated that *SNAI1* played a more important role in stimulating proliferation in HCT116 cells than in DLD1-*SNAI1* cells, with HCT116-*SNAI1* cells exhibiting an approx. 7-fold increase in proliferative capacity, whereas DLD1-*SNAI1* cells showed a more moderate increase of approx. 2-fold (Fig. 2A,B), suggesting a cell-line-dependent effect of *SNAI1* on proliferation. The effect of *SNAI1* on cell survival under stress conditions was further assessed. The survival rates of HCT116-*SNAI1* and DLD1-*SNAI1* cells were significantly increased (Fig. 2C,D), indicating that *SNAI1* significantly enhanced the tolerance of both cell lines to adverse conditions.

The results of the migration assay demonstrated that *SNAI1* overexpression markedly enhanced cell migration. Specifically, overexpression of *SNAI1* significantly increased the migratory capacity of HCT116-*SNAI1* and DLD1-*SNAI1* cells (Fig. 3A,B). To further investigate how *SNAI1* influences the phenotypic transition of tumor cells, western blot analysis was used to assess the expression of the cytoskeletal protein vimentin and the cell adhesion molecule E-cadherin. Vimentin expression in HCT116-*SNAI1* and DLD1-*SNAI1* cells was significantly upregulated, whereas E-cadherin expression was downregulated (Fig. 3C,D), suggesting that *SNAI1* overexpression promoted phenotypic transformation of tumor cells and enhanced their invasiveness.

Inhibitory effect of *SNAI1* on miR-145 expression and promoter activity

We investigated the effects of *SNAI1* overexpression on miR-145 expression in the colon cancer cell lines HCT116 and DLD1 using RT-PCR. The results showed a significant reduction in miR-145 expression in HCT116-*SNAI1* cells and a modest downregulation of miR-145 expression in DLD1-*SNAI1* cells, indicating that *SNAI1* overexpression successfully reduced miR-145 expression (Fig. 4A). Further analysis showed that the core sequence of the *SNAI1* binding site, CA(C/G)(C/G)TG, in the promoter region of miR-145 was consistent with the upstream CT-enriched region. Using luciferase reporter gene assays, we investigated how *SNAI1* affected miR-145 promoter activity (Fig. 4B). The findings demonstrated that luciferase activity was dramatically reduced in HCT116-*SNAI1* cells, suggesting that *SNAI1* strongly suppressed miR-145 promoter activity. Luciferase activity was also significantly reduced in DLD1-*SNAI1* cells, further confirming that *SNAI1* effectively inhibits miR-145 transcription.

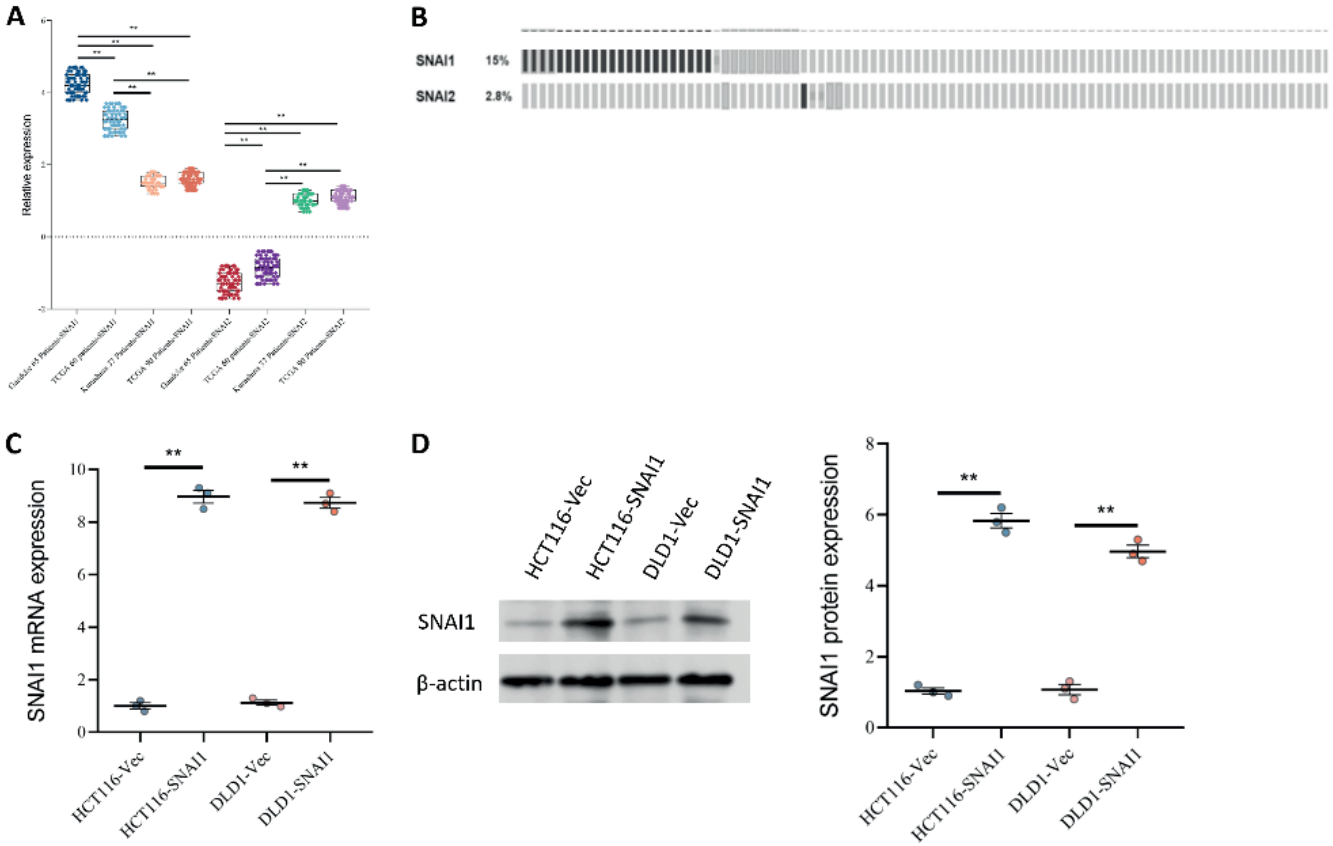


Fig. 1. Expression and genetic variation of *SNAI1* and *SNAI2* in patients with colon cancer

A. Relative expression levels of *SNAI1* and *SNAI2* in colorectal cancer (CRC) patient datasets from Kurashina (37 patients), Gaedcke (65 patients), The Cancer Genome Atlas (TCGA; 60 patients), and TCGA (90 patients), shown as fold changes relative to normal tissues; B. Frequency of *SNAI1* and *SNAI2* upregulation in CRC patients from the TCGA dataset. Percentages indicate the proportion of samples with elevated expression; C. *SNAI1* mRNA expression levels in HCT116 and DLD1 cell lines transfected with empty vector (Vec) or *SNAI1* overexpression plasmid (*SNAI1*). Data are presented as mean ± standard deviation (SD); D. Western blot analysis of *SNAI1* protein expression in HCT116 and DLD1 cell lines transfected with the *SNAI1* overexpression plasmid (*SNAI1*) or empty vector (Vec). β-actin was used as a loading control. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. In Fig. 1A, Kolmogorov–Smirnov tests (for n >50) and Shapiro–Wilk tests (for n <50) were performed to assess normality, and the Brown–Forsythe test was used to assess homogeneity of variance. Parametric analyses were performed using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test (Supplementary Tables 1–3). In Fig. 1C,D, nonparametric analyses were performed using the Mann–Whitney U test (Supplementary Table 3); *$p < 0.05$ and **$p < 0.01$ were considered statistically significant

Analysis of miR-145 expression in colorectal cancer cell lines and tissues

Using qPCR, we examined miR-145 expression levels in both healthy colon and CRC tissues. The results showed that miR-145 was considerably downregulated in cancer tissues (Fig. 5A). Additionally, we examined miR-145 expression in several colon cancer cell lines and 1 normal colon cell line. miR-145 expression was low in the SW620 cell line and moderate in the HCT8 and HT-29 cell lines, although still lower than in the normal colon cell line (Fig. 5B). Overall, miR-145 expression was consistently lower in CRC cell lines compared with normal colon cells, suggesting that miR-145 downregulation may be a common feature of colorectal tumorigenesis.

Effects of overexpression of miR-145 on invasion and migration of tumor cells

Through qPCR analysis, we confirmed that miR-145 mimics effectively increased miR-145 expression (Fig. 6A), indicating successful overexpression. The effect of miR-145 overexpression on tumor cell invasion and migration was further evaluated. The findings demonstrated a significant decrease in the invasion index of miR-145 mimic-transfected cells (Fig. 6B), suggesting that miR-145 overexpression inhibited tumor cell invasion. In addition, miR-145 overexpression markedly impaired the migratory capacity of tumor cells (Fig. 6C, Fig. 7A,B). Collectively, these results demonstrate that miR-145 acts as a negative regulator of CRC cell aggressiveness and migration.

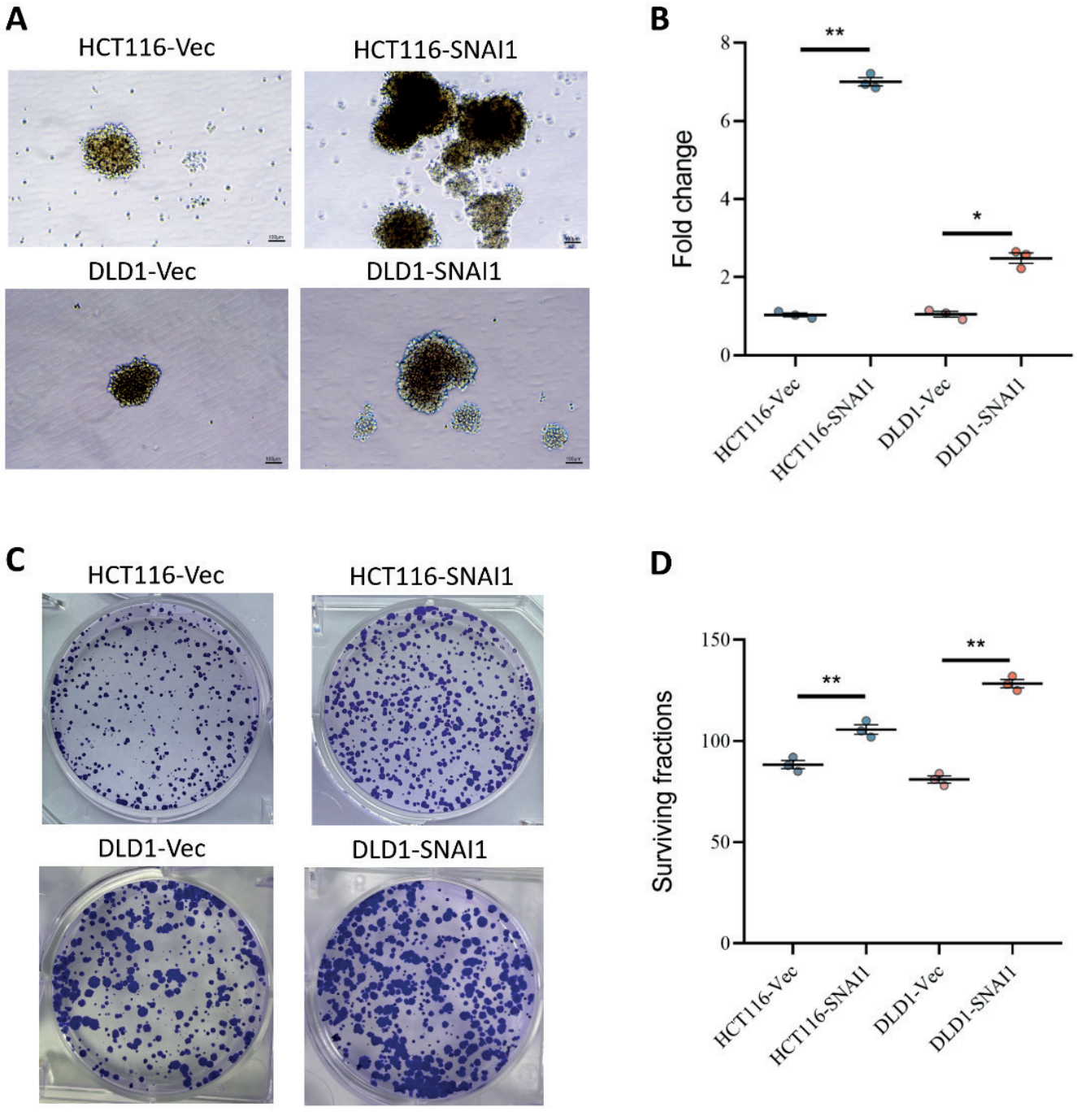


Fig. 2. Effects of *SNAI1* overexpression on cancer stemness and clonogenicity in colorectal cancer (CRC) cell lines

A,B. Sphere formation assay in HCT116 and DLD1 cell lines transfected with empty vector (Vec) or *SNAI1* overexpression plasmid (*SNAI1*); C,D. Clonogenic assay showing colony formation in HCT116 and DLD1 cell lines transfected with Vec or *SNAI1*. Representative images (left) and quantification of surviving fractions (right) are shown. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 2B,D (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

Effects of *SNAI1* on colon cancer cell survival and invasiveness and how miR-145 regulates the cAMP/PKA signaling pathway

The effects of miR-145 and *SNAI1* on DLD1 cell survival were evaluated. Overexpression of miR-145 significantly reduced cell viability, suggesting an inhibitory effect on cell survival. In contrast, overexpression of *SNAI1* resulted in a significant increase in cell viability, indicating enhanced cell survival. Notably, co-overexpression of *SNAI1* and miR-145 suggested that miR-145 could effectively counteract the survival-promoting effect of *SNAI1* (Fig. 8A,B).

In the invasion assay, overexpression of miR-145 significantly reduced the cellular invasion index, indicating

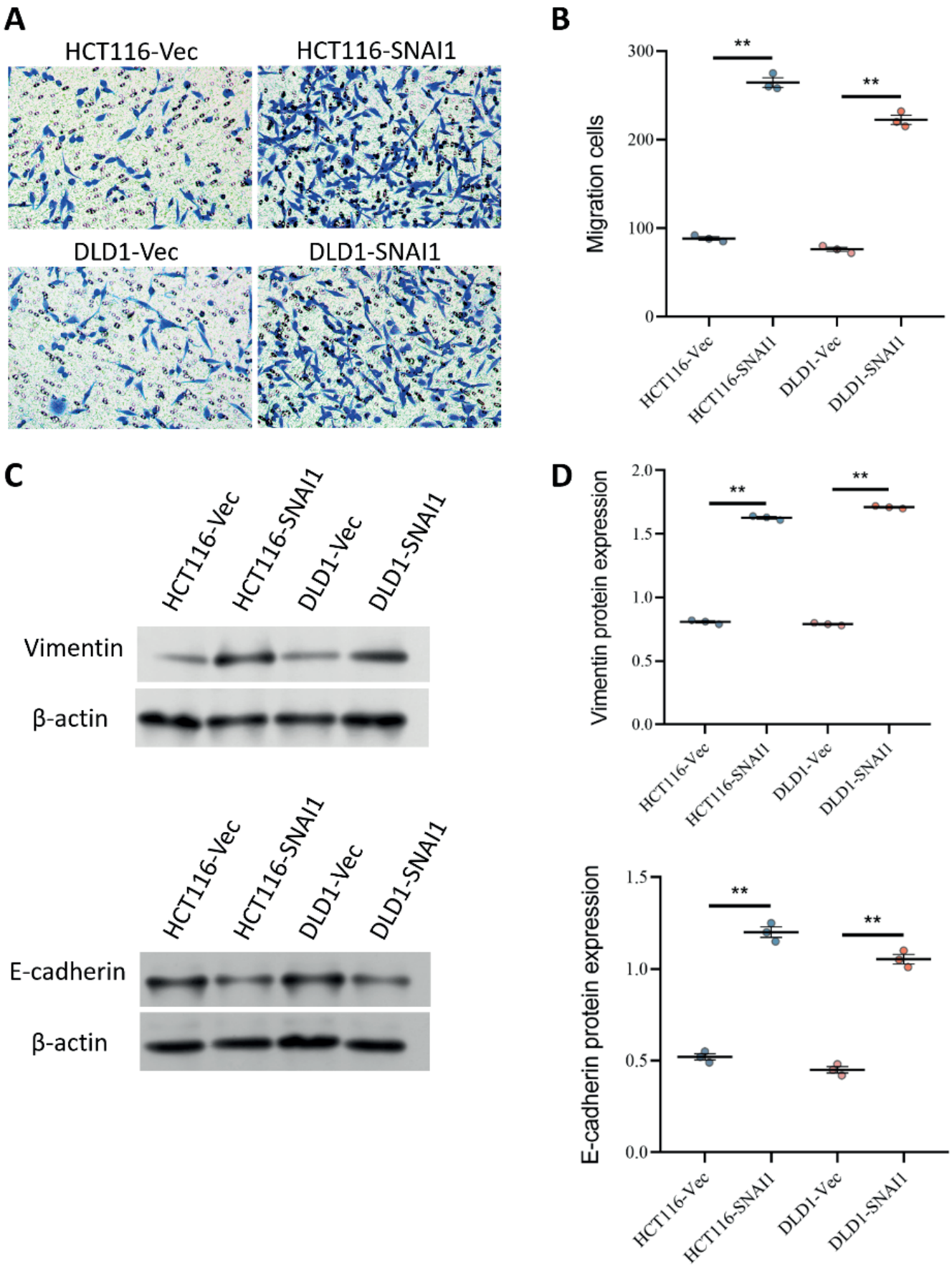


Fig. 3. Effects of *SNAI1* overexpression on migration and epithelial–mesenchymal transition (EMT)

A,B. Migration assay in HCT116 and DLD1 cell lines transfected with Vec or *SNAI1*. Representative crystal violet-stained images of migrated cells and quantification of migrated cell counts are shown; C,D. EMT markers (vimentin and E-cadherin) in HCT116 and DLD1 cell lines transfected with Vec or *SNAI1* were examined using western blot analysis. β-actin was used as a loading control. Protein expression normalized to β-actin is presented as mean ± standard deviation (SD) in bar graphs. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 3B,D (Supplementary Table 3); *$p < 0.05$ and **$p < 0.01$ were considered statistically significant

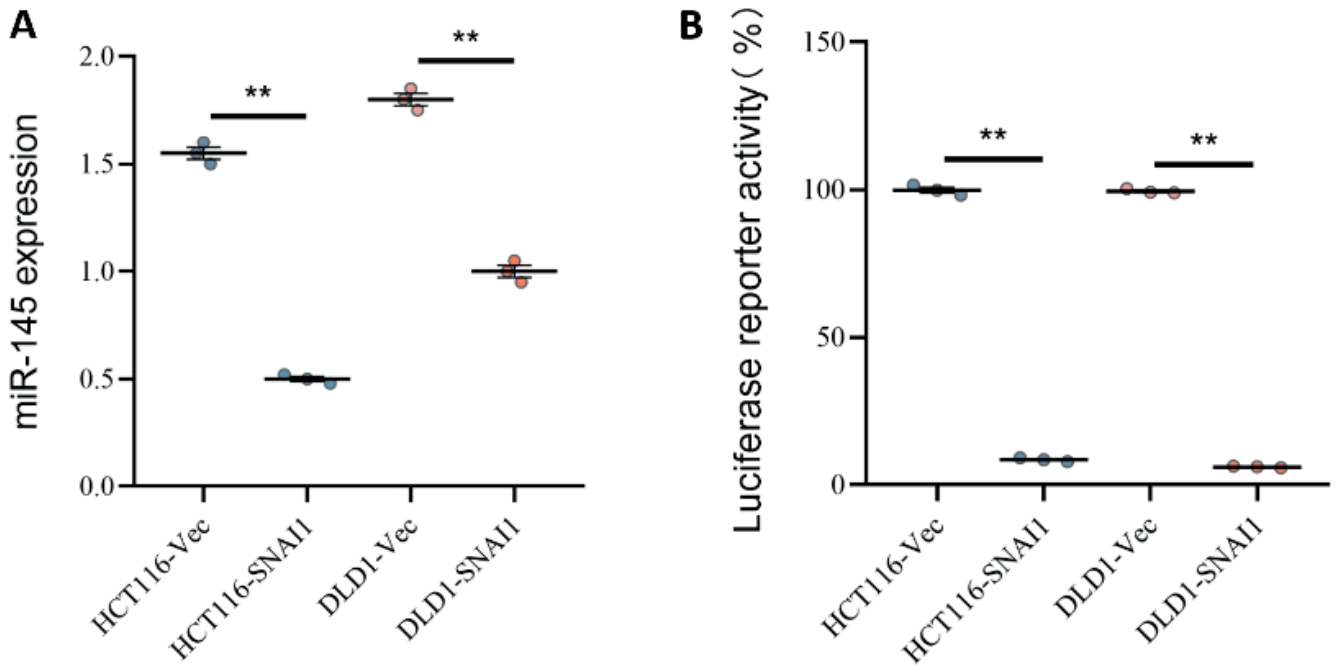


Fig. 4. Impact of *SNAI1* overexpression on miR-145 expression and target reporter activity in colorectal cancer (CRC) cell lines

A. Relative expression levels of miR-145 in HCT116 and DLD1 cell lines transfected with empty vector (Vec) or *SNAI1* overexpression plasmid (*SNAI1*); B. Luciferase reporter assay showing miR-145 target reporter activity in HCT116 and DLD1 cell lines transfected with Vec or *SNAI1*. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 4A,B (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

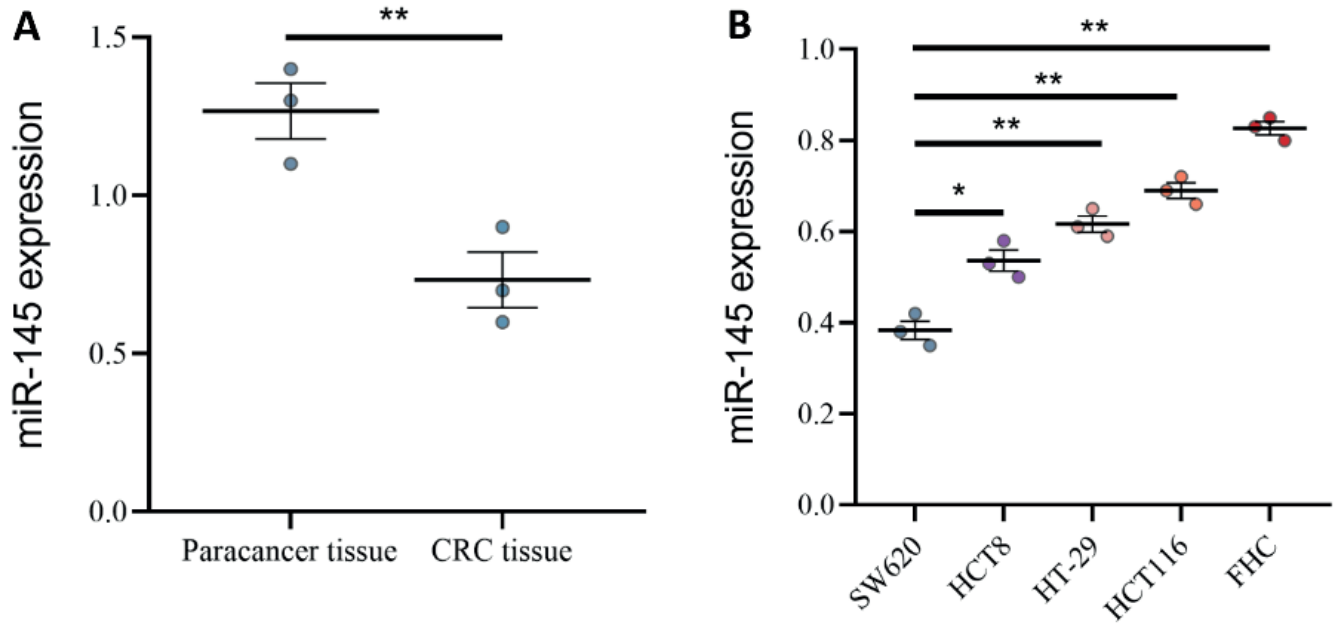


Fig. 5. Expression of miR-145 in colorectal cancer (CRC) tissues and cell lines

A. MiR-145 expression levels in paracancerous and CRC tissues; B. Relative expression levels of miR-145 in different CRC cell lines (SW620, HCT8, HT-29, and HCT116) compared with normal colonic epithelial cells (FHC). Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. In Fig. 5A, nonparametric analyses were performed using the Mann–Whitney U test. In Fig. 5B, nonparametric analyses were performed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction; the number of corrections corresponded to the total number of pairwise group comparisons (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

the ability of miR-145 to inhibit cell invasion. In contrast, overexpression of *SNAI1* enhanced the invasive capacity of the cells. Notably, when both *SNAI1* and miR-145 were overexpressed, the results suggested that miR-145 inhibited tumor cell invasion even in the presence of *SNAI1* overexpression (Fig. 9A,B).

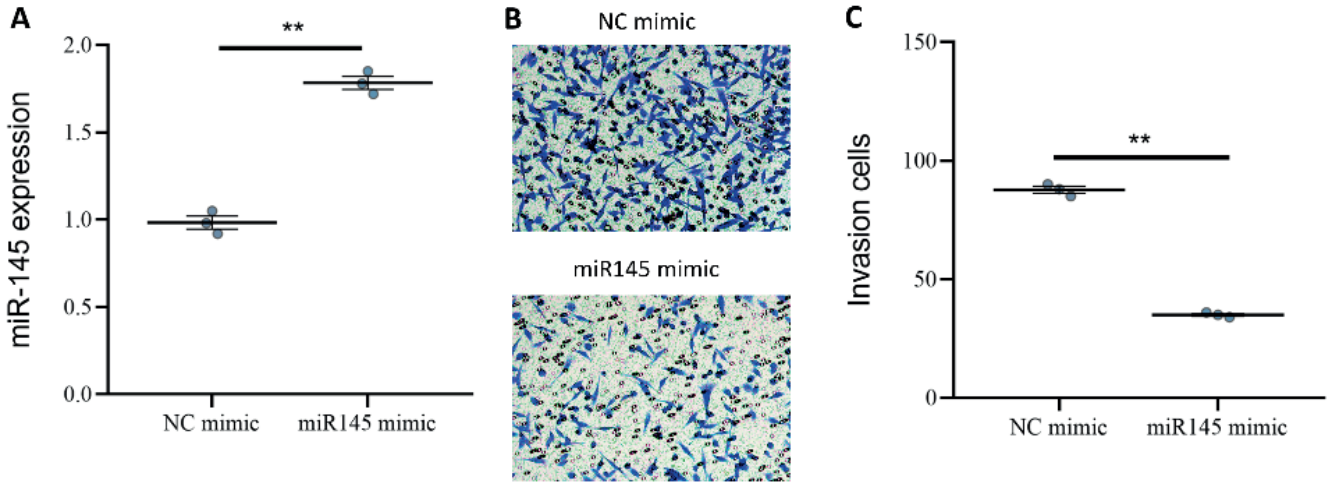


Fig. 6. Effects of miR-145 mimic on colorectal cancer (CRC) cell invasion

A. Quantitative real-time PCR (qPCR) analysis of miR-145 expression in CRC cells transfected with either miR-145 mimic or negative control (NC mimic); B,C. Transwell invasion assay of CRC cells transfected with NC mimic or miR-145 mimic. Quantification of invading cells and representative images are shown. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 6A,B (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

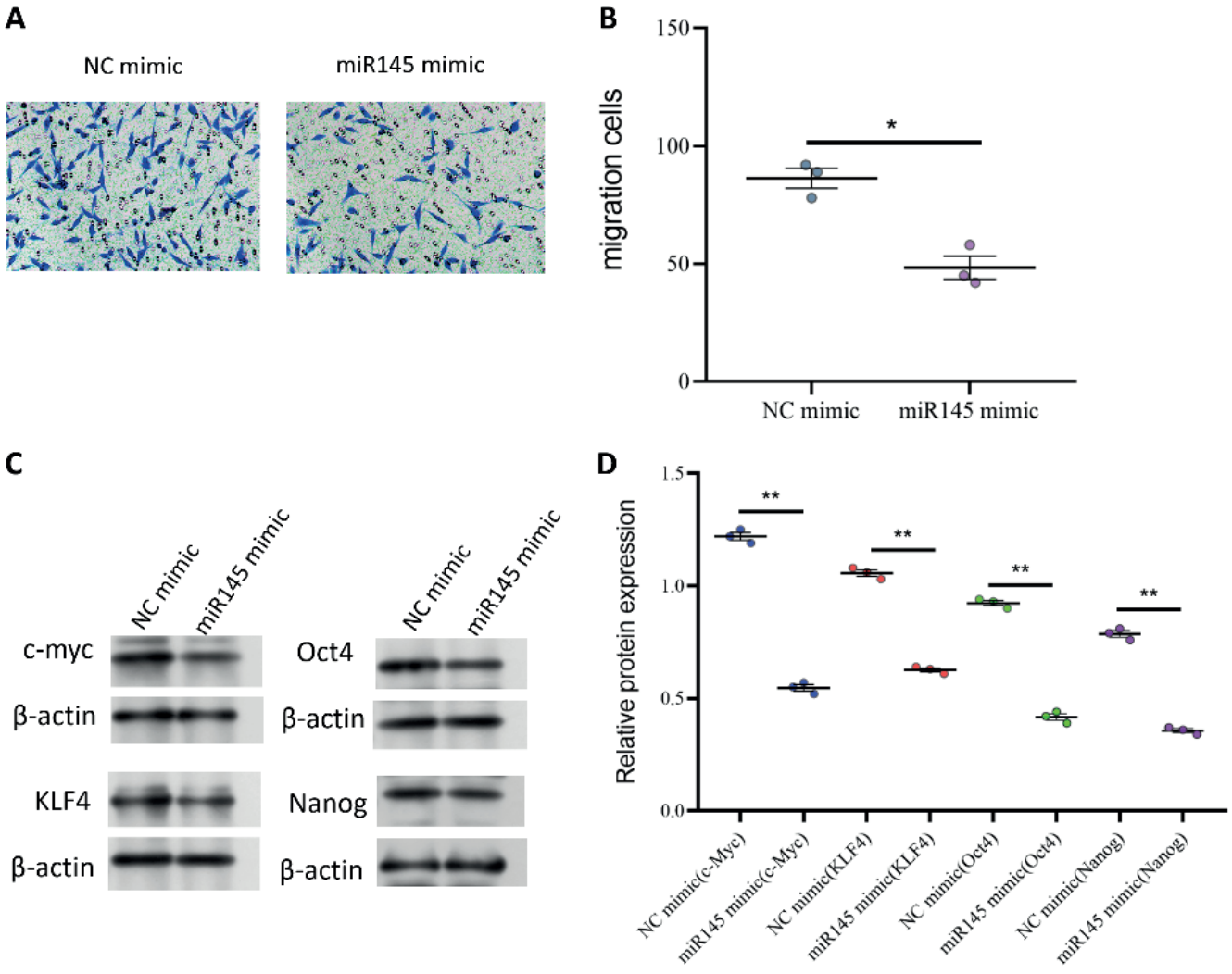


Fig. 7. Effects of miR-145 mimic on colorectal cancer (CRC) cell migration and stemness marker expression

A,B. Migration assay of CRC cells transfected with negative control (NC) mimic or miR-145 mimic. Quantification of migratory cells and representative images are shown; C,D. Western blot analysis of CRC cells transfected with miR-145 mimic or NC mimic using stemness-associated markers (c-Myc, KLF4, Oct4, and Nanog). β-actin was used as a loading control. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 7A,B (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

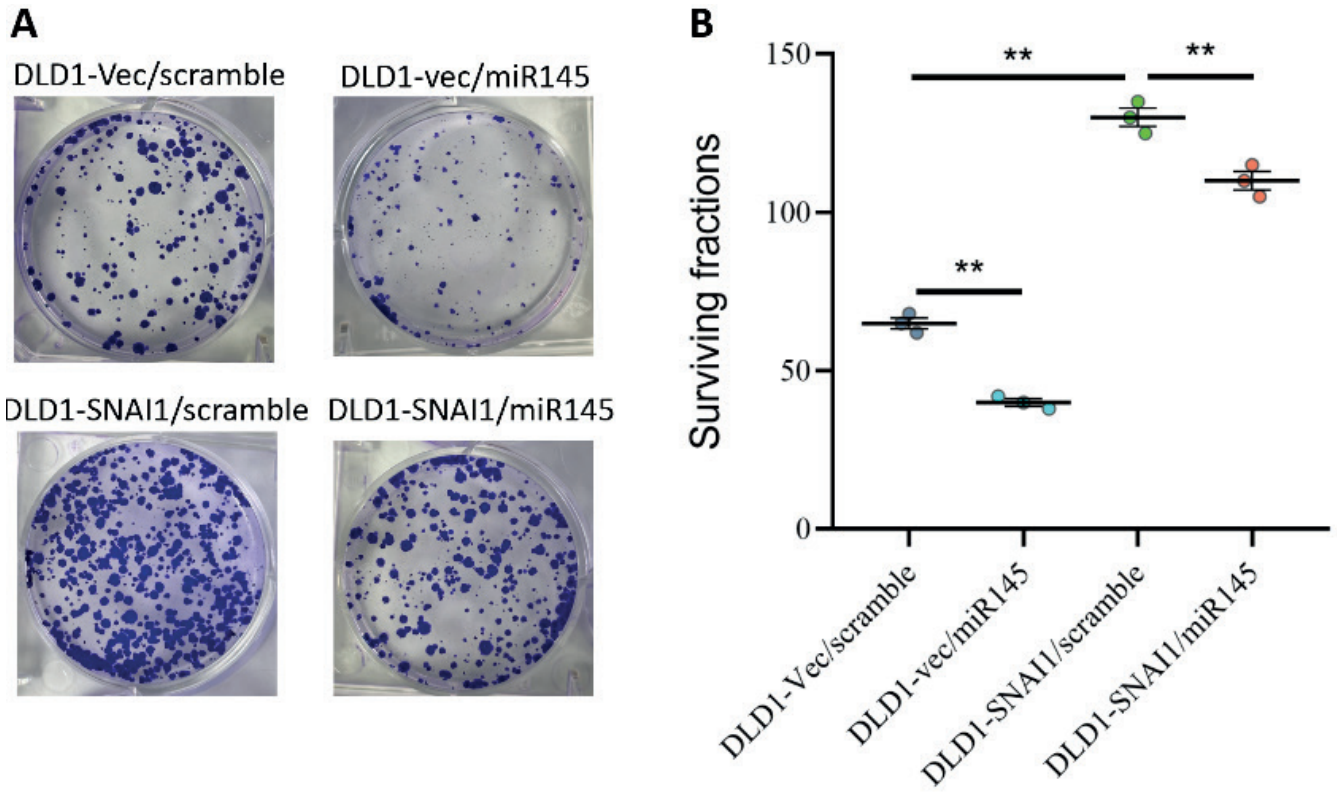


Fig. 8. Effects of miR-145 and *SNAI1* expression on clonogenicity in DLD1 cells

Clonogenic assay in DLD1 cells treated with miR-145 mimic or scramble control following transfection with either vector control (Vec) or *SNAI1* overexpression plasmid (*SNAI1*). Representative images of colony formation and quantification of surviving fractions are shown. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. In Fig. 8B, nonparametric analyses were performed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

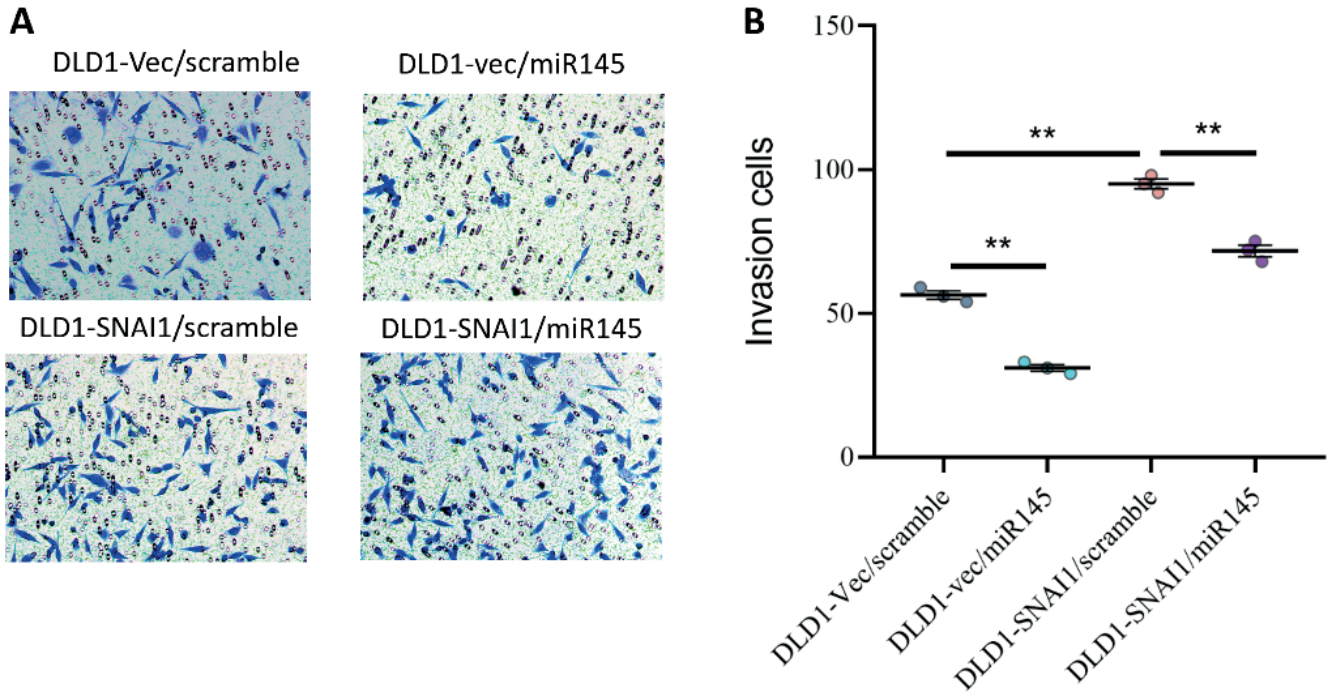


Fig. 9. Effects of miR-145 and *SNAI1* expression on invasion in DLD1 cells

Transwell invasion assay in DLD1 cells transfected and treated as described in panel A. Representative images of invading cells and quantitative analysis are shown. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. In Fig. 9B, nonparametric analyses were performed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

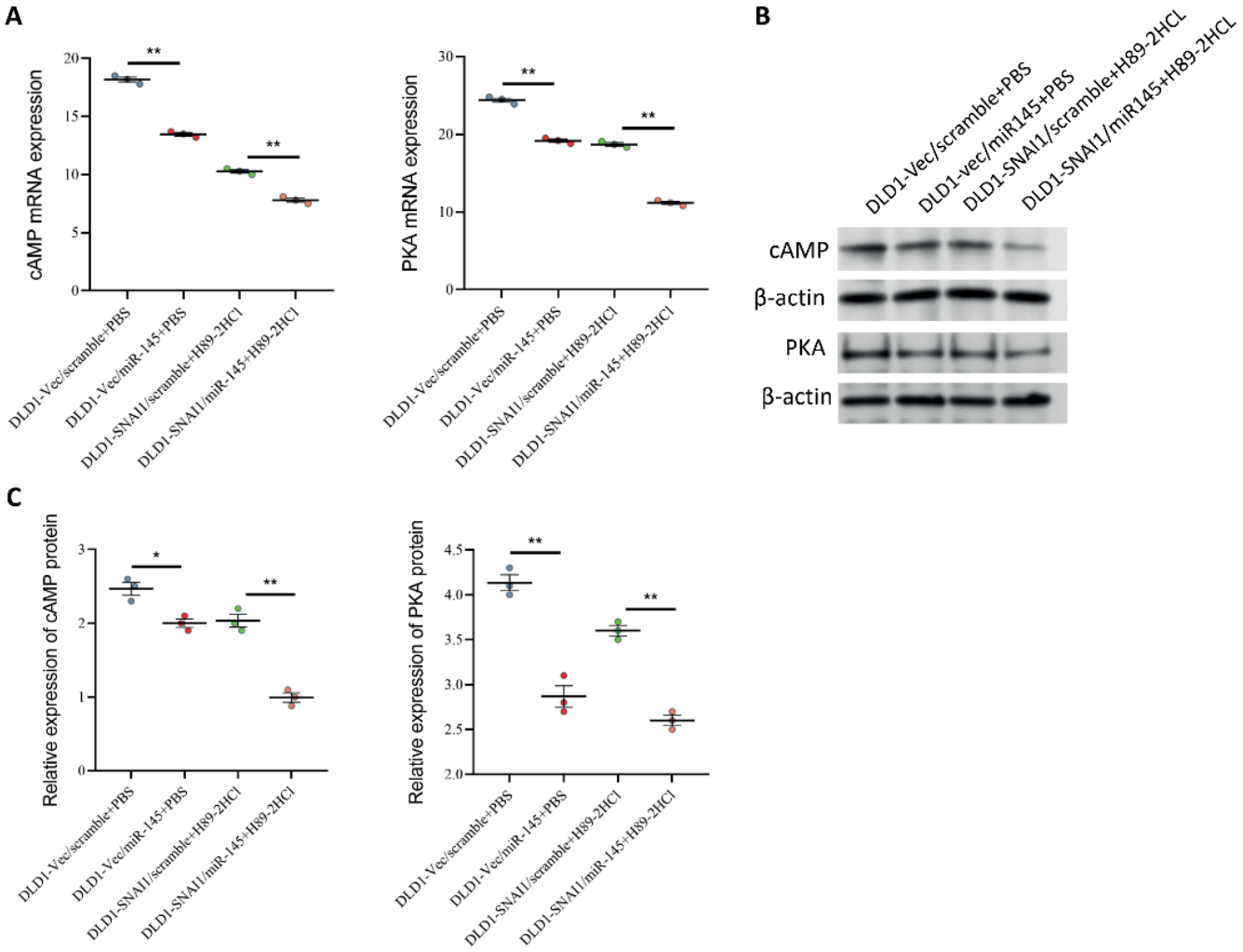


Fig. 10. Effects of miR-145 and *SNAI1* expression on cAMP/PKA signaling in DLD1 cells

A. Comparative cAMP and PKA expression levels in DLD1 cells treated with miR-145 mimic or scramble control, with or without the PKA inhibitor H89-2HCl; B,C. Western blot analysis of cAMP and PKA protein levels in DLD1 cells. Data are presented as dots, with horizontal lines indicating medians. Horizontal lines above the dots and asterisks (*) indicate statistically significant differences between the 2 connected groups. Nonparametric analyses were performed using the Mann–Whitney U test in Fig. 10A,B (Supplementary Table 3); *p < 0.05 and **p < 0.01 were considered statistically significant

Subsequent experiments revealed that miR-145 overexpression markedly decreased cAMP and PKA activity, suggesting that miR-145 may exert its effects by blocking the cAMP/PKA pathway. Use of the PKA inhibitor H89 also significantly reduced cAMP and PKA levels, indicating effective inhibition of PKA activity (Fig. 10A). In the context of miR-145 overexpression, H89 further reduced cAMP and PKA levels, suggesting that the 2 may synergistically inhibit activity of the cAMP/PKA pathway (Fig. 10B,C).

Discussion

A number of hereditary and environmental factors are closely associated with the incidence of CRC, one of the malignant tumors with the highest morbidity and mortality rates worldwide. According to global cancer statistics, colon cancer ranked 3rd in terms of incidence and 2nd in terms of mortality among malignant tumors in 2020, with approx. 1.88 million new cases and 920,000 deaths.[10] The incidence and mortality of CRC continue to rise in developing countries, particularly China, despite improvements in early detection and treatment resulting from advances in screening technologies, especially in industrialized nations.[11] Therefore, CRC prevention and treatment remain major global public health challenges.[12]

The development and progression of CRC involve a complex, multistep process characterized by interactions among numerous genes and signaling pathways.[13,14] In recent years, increasing attention has been directed toward the role of EMT in tumor invasion, metastasis, and chemoresistance. One of the key steps enabling tumor dissemination is the transition of epithelial cells into more aggressive and migratory mesenchymal cells through a series of molecular alterations, known as EMT.[15,16] In CRC, Snail family transcription factors such as *SNAI1* and *SNAI2* are considered core regulators of the EMT process, promoting

tumor cell migration and invasion through regulation of cytoskeletal remodeling, adhesion molecule expression, and cell–cell interactions.[17,18]

In this study, we focused on *SNAI1* and *SNAI2*, 2 key transcription factors, and their roles in CRC. We compared the expression patterns of *SNAI1* and *SNAI2* among different cohorts of colon cancer patients and further analyzed their genetic alterations. Through comparative analysis of multiple datasets, we found that *SNAI1* was significantly overexpressed in certain colon cancer patients, whereas *SNAI2* was suppressed in others.[19] Furthermore, using cell-based experiments, we investigated the role of *SNAI1* in colon cancer cell migration, proliferation, phenotypic transformation, and interaction with miR-145.[20]

The study found that *SNAI1* not only promoted proliferation and survival of colon cancer cells, but also significantly enhanced migratory capacity. In particular, overexpression of *SNAI1* in HCT116 cells increased proliferative capacity approx. 7-fold, and the survival rate was also significantly increased. In addition, *SNAI1* overexpression promoted phenotypic transformation of the cells, as demonstrated by increased expression of the cytoskeletal protein vimentin and decreased expression of the cell adhesion molecule E-cadherin, thereby further enhancing cellular aggressiveness.[21]

We investigated the molecular interaction between *SNAI1* and miR-145. MiR-145, a well-known tumor suppressor, has been shown to inhibit tumor growth and metastasis in a variety of tumor types.[22,23] In the present study, *SNAI1* overexpression markedly reduced miR-145 expression. These findings suggest that *SNAI1* may regulate miR-145 expression at the transcriptional level by modulating miR-145 promoter activity. Furthermore, we found that miR-145 expression was frequently downregulated in colon cancer cell lines and tissues, and that its overexpression inhibited tumor cell invasion and migration, further supporting the role of miR-145 as a putative tumor suppressor.[24]

Our findings suggest that *SNAI1* and miR-145 play pivotal roles in modulating key biological processes driving CRC progression. *SNAI1*, a well-known transcription factor promoting EMT, facilitates tumor cell invasion, metastasis, and therapeutic resistance. In contrast, miR-145 functions as a tumor suppressor by inhibiting pathways associated with proliferation and migration, partly through targeting *SNAI1* and regulating the cAMP/PKA signaling axis. Dysregulation of the *SNAI1*/miR-145 axis may therefore enhance tumor aggressiveness and may be associated with poor patient prognosis. These findings underscore the significance of *SNAI1* and miR-145 not only in cellular behavior but also in the broader biological processes underlying CRC progression. By suppressing miR-145 expression, *SNAI1* may not only serve as a crucial regulatory component in the EMT process of CRC cells but may also contribute to increased tumor aggressiveness and metastasis. Furthermore, upregulation of miR-145 may effectively reduce tumor cell aggressiveness.

While our study focused on the regulatory effects of *SNAI1* and miR-145 on the cAMP/PKA pathway, their potential interactions with other critical signaling cascades, such as the Wnt/β-catenin, PI3K/AKT, and TGF-β pathways, remain to be elucidated. Future studies are needed to explore these associations in order to fully understand the broader regulatory network and oncogenic potential of *SNAI1* and miR-145 in CRC.

Potential clinical implications

Given the pivotal role of the *SNAI1*/miR-145 regulatory axis in CRC progression, targeting this pathway represents a promising therapeutic strategy. Potential approaches include development of miR-145 mimics to restore its tumor-suppressive function, use of small-molecule inhibitors to suppress *SNAI1* expression or activity, and development of gene-therapy approaches to modulate the expression levels of these key molecules. However, clinical application of these strategies requires careful consideration of delivery efficiency, target specificity, and potential off-target effects. Further preclinical studies and clinical trials are needed to evaluate the safety and efficacy of these approaches in patients with CRC.

Limitations of the study

This study has several limitations that should be acknowledged. First, our findings are primarily based on in vitro experiments using 2 CRC cell lines (DLD1 and HCT116), which may not fully reflect the complexity and heterogeneity of CRC. Validation using additional cell lines with diverse genetic backgrounds is needed to strengthen the generalizability of our conclusions. Second, the lack of in vivo studies limits our ability to fully elucidate the roles of *SNAI1* and miR-145 within the physiological tumor microenvironment (TME), where factors such as blood supply, immune surveillance, and stromal interactions play critical roles. Future investigations using animal models or patient-derived xenografts (PDXs) are warranted. Third, although we demonstrated that *SNAI1* and miR-145 regulate the cAMP/PKA signaling pathway, their interactions with other key oncogenic pathways, such as Wnt/β-catenin, PI3K/AKT, and TGF-β, remain to be explored. Finally, the clinical relevance of our findings needs to be verified through analysis of patient samples to assess the potential of *SNAI1* and miR-145 as diagnostic biomarkers or therapeutic targets.

Moreover, technical factors such as variations in transfection efficiency, reagent handling, and measurement precision may introduce bias. Potential off-target effects and batch-to-batch variability of experimental reagents could also affect the results. Rigorous experimental controls and repeated validation experiments will be necessary to confirm the reproducibility of our findings.

Conclusions

This work demonstrated the critical roles of *SNAI1* and miR-145 in regulating the cAMP/PKA signaling pathway, as well as colon cancer cell survival and invasiveness. The findings further supported the critical role of *SNAI1* in tumor progression by demonstrating that *SNAI1* overexpression enhanced tumor cell survival and invasive potential through upregulation of cAMP and PKA levels. In contrast, overexpression of miR-145 showed promise as a tumor suppressor by markedly inhibiting the cAMP/PKA pathway and reducing tumor cell invasiveness and survival.

When *SNAI1* and miR-145 were simultaneously overexpressed, miR-145 counteracted the tumor-promoting effects of *SNAI1* by significantly downregulating cAMP/PKA activity, suggesting that miR-145 may antagonize *SNAI1*-mediated tumor promotion. In addition, administration of the PKA inhibitor H89 further attenuated *SNAI1*-induced tumor progression, confirming the important role of the cAMP/PKA pathway in *SNAI1* regulation. These results provide novel insights into the therapeutic potential of miR-145 and *SNAI1* as targets for cancer treatment.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15684050. The package contains the following files:

Supplementary Table 1. The results of normality test.

Supplementary Table 2. The results of homogeneity of variance.

Supplementary Table 3. The results of statistical analysis.

Data Availability Statement

The data that support the findings of this study are openly available in Figshare at https://figshare.com/s/513f4d57719ed02098c4.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Jianshan Cai https://orcid.org/0009-0002-9133-1915
Qiang Sun https://orcid.org/0009-0008-3623-8041
Shichao Deng https://orcid.org/0009-0003-4411-6857
Qi Wei https://orcid.org/0009-0005-3124-6550
Longzhi Li https://orcid.org/0009-0004-8154-7266
Baojin Ma https://orcid.org/0009-0001-8092-6844
Jiadong Chen https://orcid.org/0009-0002-4758-8262

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
2. Wilson BJ, Schatton T, Frank MH, Frank NY. Colorectal cancer stem cells: Biology and therapeutic implications. *Curr Colorectal Cancer Rep*. 2011;7(2):128–135. doi:10.1007/s11888-011-0093-2
3. Wu J, Li W, Su J, et al. Integration of single-cell sequencing and bulk RNA-seq to identify and develop a prognostic signature related to colorectal cancer stem cells. *Sci Rep*. 2024;14(1):12270. doi:10.1038/s41598-024-62913-3
4. Ren Y, Mao X, Xu H, et al. Ferroptosis and EMT: Key targets for combating cancer progression and therapy resistance. *Cell Mol Life Sci*. 2023;80(9):263. doi:10.1007/s00018-023-04907-4
5. Cano A, Pérez-Moreno MA, Rodrigo I, et al. The transcription factor Snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol*. 2000;2(2):76–83. doi:10.1038/35000025
6. Mozammel N, Amini M, Baradaran B, Mahdavi SZB, Hosseini SS, Mokhtarzadeh A. The function of miR-145 in colorectal cancer progression: An updated review on related signaling pathways. *Pathol Res Pract*. 2023;242:154290. doi:10.1016/j.prp.2022.154290
7. Shen X, Jiang H, Chen Z, et al. MicroRNA-145 inhibits cell migration and invasion in colorectal cancer by targeting TWIST. *Onco Targets Ther*. 2019;12:10799–10809. doi:10.2147/OTT.S216147
8. Feng Y, Zhu J, Ou C, et al. MicroRNA-145 inhibits tumour growth and metastasis in colorectal cancer by targeting fascin-1. *Br J Cancer*. 2014;110(9):2300–2309. doi:10.1038/bjc.2014.122
9. Sargolzaei J, Etemadi T, Alyasin A. The P53/microRNA network: A potential tumor suppressor with a role in anticancer therapy. *Pharmacol Res*. 2020;160:105179. doi:10.1016/j.phrs.2020.105179
10. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–263. doi:10.3322/caac.21834
11. Yang Y, Han Z, Li X, et al. Epidemiology and risk factors of colorectal cancer in China. *Chin J Cancer Res*. 2020;32(6):729–741. doi:10.21147/j.issn.1000-9604.2020.06.06
12. Li Q, Wu H, Cao M, et al. Colorectal cancer burden, trends and risk factors in China: A review and comparison with the United States. *Chin J Cancer Res*. 2022;34(5):483–495. doi:10.21147/j.issn.1000-9604.2022.05.08
13. Armaghany T, Wilson JD, Chu Q, Mills G. Genetic alterations in colorectal cancer. *Gastrointest Cancer Res*. 2012;5(1):19–27. PMID:22574233. PMCID:PMC3348713.
14. Malki A, El-Ruz RA, Gupta I, Allouch A, Vranic S, Al Moustafa AE. Molecular mechanisms of colon cancer progression and metastasis: Recent insights and advancements. *Int J Mol Sci*. 2020;22(1):130. doi:10.3390/ijms22010130
15. Vu T, Datta P. Regulation of EMT in Colorectal cancer: A culprit in metastasis. *Cancers (Basel)*. 2017;9(12):171. doi:10.3390/cancers9120171
16. Cao H, Xu E, Liu H, Wan L, Lai M. Epithelial–mesenchymal transition in colorectal cancer metastasis: A system review. *Pathol Res Pract*. 2015;211(8):557–569. doi:10.1016/j.prp.2015.05.010
17. Brzozowa M, Michalski M, Wyrobiec G, et al. The role of Snail1 transcription factor in colorectal cancer progression and metastasis. *Wspolczesna Onkol*. 2015;4:265–270. doi:10.5114/wo.2014.42173
18. Zhang N, Ng AS, Cai S, Li Q, Yang L, Kerr D. Novel therapeutic strategies: targeting epithelial–mesenchymal transition in colorectal cancer. *Lancet Oncol*. 2021;22(8):e358–e368. doi:10.1016/S1470-2045(21)00343-0
19. Hong KS, Ryu KJ, Kim H, et al. MSK1 promotes colorectal cancer metastasis by increasing Snail protein stability through USP5-mediated Snail deubiquitination. *Exp Mol Med*. 2025;57(4):820–835. doi:10.1038/s12276-025-01433-0
20. Zhu LF, Hu Y, Yang CC, et al. Snail overexpression induces an epithelial to mesenchymal transition and cancer stem cell-like properties in SCC9 cells. *Lab Invest*. 2012;92(5):744–752. doi:10.1038/labinvest.2012.8

21. Cheng X, Shen T, Liu P, et al. mir-145-5p is a suppressor of colorectal cancer at early stage, while promotes colorectal cancer metastasis at late stage through regulating AKT signaling evoked EMT-mediated anoikis. *BMC Cancer*. 2022;22(1):1151. doi:10.1186/s12885-022-10182-6
22. Zeinali T, Mansoori B, Mohammadi A, Baradaran B. Regulatory mechanisms of miR-145 expression and the importance of its function in cancer metastasis. *Biomed Pharmacother*. 2019;109:195–207. doi:10.1016/j.biopha.2018.10.037
23. Angius A, Uva P, Pira G, et al. Integrated analysis of miRNA and mRNA endorses a twenty miRNAs signature for colorectal carcinoma. *Int J Mol Sci*. 2019;20(16):4067. doi:10.3390/ijms20164067
24. Xu WX, Liu Z, Deng F, et al. MiR-145: A potential biomarker of cancer migration and invasion. *Am J Transl Res*. 2019;11(11):6739–6753. PMID:31814885. PMCID:PMC6895535.

Updating guidelines for the diagnosis of fetal alcohol spectrum disorders (FASD) in Poland

Katarzyna A. Dyląg[1,2,A–F], Małgorzata Klecka[3,4,C–F], Magdalena Borkowska[5,A–D,F], Iwona Palicka[3,6,B–F], Katarzyna Okulicz-Kozaryn[7,C–F], Agata Cichoń-Chojnacka[8,9,C–F], Tomasz Maciejewski[10,C,E,F], Robert S. Śmigiel[11,A,–F]

[1] Department of Pathophysiology, Jagiellonian University Medical College, Kraków, Poland
[2] St. Louis Children Hospital, Kraków, Poland
[3] Fastryga Foundation, Świerklany, Poland
[4] Diagnostic and Therapeutic Center for Developmental Disorders, Świerklany, Poland
[5] FASada Diagnostic and Therapy Centre, Warszawa, Poland
[6] Child Development Stimulation Centre, Poznań, Poland
[7] Department of Children and Adolescents Health, Institute of Mother and Child, Warszawa, Poland
[8] Centre for FASD Diagnosis, Gdynia, Poland
[9] Child Psychiatric Service KEJA-MED, Gdynia, Poland
[10] Department of Health Sociology, Education and Medical Communication, Institute of Mother and Child, Warszawa, Poland
[11] Department of Pediatrics, Endocrinology, Diabetology and Metabolic Diseases, Wroclaw Medical University, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1063–1072

Address for correspondence
Katarzyna Okulicz-Kozaryn
E-mail: katarzyna.okulicz@imid.med.pl

Funding sources
None declared

Conflict of interest
None declared

Received on October 31, 2025
Reviewed on November 30, 2025
Accepted on January 28, 2026

Published online on March 5, 2026

Cite as
Dyląg KA, Klecka M, Borkowska M, et al. Updating guidelines for the diagnosis of fetal alcohol spectrum disorders (FASD) in Poland. *Adv Clin Exp Med*. 2026;35(6):1063–1072. doi:10.17219/acem/217484

DOI
10.17219/acem/217484

Abstract

Background. The prevalence of fetal alcohol spectrum disorders (FASD) in Poland is about 2%, but the prevalence of prenatal alcohol exposure (PAE) is much higher than previously reported. In the absence of a single biomarker or imaging test that can confirm the diagnosis, the identification of FASD relies on clinical diagnostic criteria; therefore, the first Polish diagnostic standards were developed in 2020.

Objectives. To present the process of updating the national diagnostic guidelines for FASD in Poland and discuss the key revisions made based on clinical implementation feedback.

Materials and methods. The work was carried out by the members of the Council for FASD Prevention and Treatment at the National Centre for Addiction Prevention in 4 steps: 1) a pilot implementation study; 2) a structured literature review supporting the update; 3) a formalized expert consensus process; and 4) stakeholder consultation.

Results. The updated guidelines include the following sections: Introduction, Methodology, Diagnostic categories, Diagnostic scheme, Ethical considerations, Evaluation of PAE, Evaluation of facial dysmorphism, Evaluation of growth impairment, and Evaluation of the central nervous system (CNS). Appendices containing practical tools useful in the diagnostic procedure are an important element of the recommendations. Key amendments include the introduction of the partial fetal alcohol syndrome (pFAS) diagnostic category; the addition of biomarker analysis as a tool to confirm PAE; simplification of the assessment of facial dysmorphologies; clarification of CNS evaluation; and discussion of the ethical concerns associated with FASD diagnosis.

Conclusions. The updated national guidelines may improve the quality and standardization of FASD diagnosis not only in Poland, but also worldwide. The practical utility of each recommendation should be continuously monitored, validated, and updated.

Key words: Poland, diagnostic guidelines, fetal alcohol spectrum disorders (FASD)

Acknowledgements

The authors of this report would like to acknowledge the National Centre for Prevention of Addictions – the institution that coordinated and sponsored the development of the Polish 2025 guidelines. The board members (Bogusława Bukowska, Katarzyna Łukowska), who supported the process, as well as other individuals (Bartosz Kehl, Sylwia Opasińska, Aneta Stankowska) who provided help and assistance, are gratefully acknowledged. Special thanks are also due to the other contributors to the 2025 Polish guidelines – Teresa Jadczak-Szumiło, Katarzyna Kałamajska-Liszcz, Katarzyna Przybyszewska, Krystyna Szymańska, Seweryna Konieczna, Iwona Sawionek, Jolanta Terlikowska, and Andrzej Urbanik – as well as to the authors of the 2020 guidelines.

Highlights

- Updated Polish national guidelines standardize fetal alcohol spectrum disorder (FASD) diagnosis, addressing the high prevalence of prenatal alcohol exposure and reducing diagnostic variability.
- The introduction of partial fetal alcohol syndrome (pFAS) and the use of biomarkers enhance assessment of prenatal alcohol exposure, improving diagnostic accuracy in the absence of a single definitive test.
- Revised criteria simplify the evaluation of facial dysmorphology and clarify central nervous system assessment, improving clinical feasibility and inter-rater consistency.
- These evidence-based, consensus-driven FASD guidelines, supported by practical diagnostic tools, may improve global clinical practice by promoting ethical, reliable, and reproducible diagnosis.

Background

Fetal alcohol spectrum disorders (FASD) constitute an umbrella term describing a range of conditions resulting from prenatal alcohol exposure (PAE). The global prevalence of FASD is estimated at 22.77 per 1,000 individuals (range: 0–176.77 per 1,000).[1,2] The most recent study attempting to assess the prevalence of FASD among Polish children was conducted in 2013 and indicated that at least 2% of children in Poland are affected.[3] However, more recent findings suggest that alcohol consumption during pregnancy may be considerably more common than previously reported.[4,5] Despite increasing awareness, FASD remains a significant public health concern both in Poland[4,6] and globally.[7]

In the absence of a single biomarker or imaging modality capable of confirming the diagnosis,[8,9] the identification of FASD – unlike many other medical conditions – relies primarily on clinical diagnostic criteria rather than laboratory findings. Consequently, standardized diagnostic guidelines remain the gold standard for diagnosing FASD. Since the first description of the dysmorphic features of fetal alcohol syndrome (FAS) by Jones and Smith,[10] multiple diagnostic systems for FASD have been developed.

In North America, 4 principal diagnostic systems are currently used: the Institute of Medicine (IOM)/Hoyme et al. criteria,[11] the 4-Digit Diagnostic Code,[12] the Emory criteria,[13] and the Canadian guidelines.[14] However, inconsistencies among these systems have been documented.[15,16] Recently, the need for greater standardization of FASD diagnostic guidelines has been strongly emphasized.[16] Over the past several years, a number of countries – including Germany,[17,18] Norway,[19] Italy,[20] and Australia[21] – have developed their own national diagnostic guidelines. The first Polish diagnostic standards were developed in 2020.[22] In 2023, the National Centre for Prevention of Addictions conducted a nationwide survey to assess their applicability in Polish FASD diagnostic centers.[23] The results indicated that the standards were known in detail by about half (58%) of the study participants, mainly because their dissemination was limited (according to 90% of respondents). More detailed results of the survey indicated that the respondents agreed regarding the general value of the document: 100% of them rated the guidelines as practical, useful, and well structured.

The diagnostic scheme was positively assessed by 95% of the survey participants. Respondents' answers differed regarding diagnostic categories (47% rated them negatively, 52% positively, and 1% left the question unanswered). Common reasons for negative assessments included the view that neurodevelopmental disorder associated with prenatal alcohol exposure (ND-PAE) constitutes overly broad criteria, the need for clearer differentiation between FASD diagnostic categories, and the lack of pFAS as a diagnostic category.

The part of the guidelines concerning PAE assessment (including the Alcohol Use Disorder Identification Test – Consumption (AUDIT-C) and Alcohol Use Disorder Identification Test (AUDIT) as part of it) was rated positively by 68% of the respondents, while the remaining 32% rated it negatively. Common reasons for negative assessments included the statement that biological parents often cannot be reached and that legal guardians rarely have knowledge about PAE; moreover, the diagnosis of PAE requires the ability to conduct a therapeutic conversation due to the mother's feelings of guilt and shame.

Instructions on the physical examination, particularly in terms of the proposed tools (growth charts), were viewed positively by 89% of the participants. The proposed neuropsychological criteria were approved by 90% of the participants; however, 26% considered the neuropsychological criteria inadequate for young children. These results confirm the need to update and disseminate the FASD diagnostic guidelines among Polish specialists in order to improve healthcare for individuals affected by PAE.

Objectives

This paper summarizes the process of updating the national diagnostic guidelines for FASD in Poland and discusses key revisions made based on clinical implementation feedback. The main objective of this work was to improve the quality of FASD diagnosis and provide solid background for healthcare of individuals affected by PAE. The specific objectives concerned the following stages of work:

1. To evaluate functioning of the first guidelines (from 2020) in clinical practice;
2. To update the knowledge on FASD;
3. To refine and synthesize experts' opinions and gain consensus on the guidelines update;
4. Stakeholder consultation.

In this paper, we briefly present these 4 stages of the updating process and an abbreviated version of the recently updated Polish diagnostic guidelines already published in the Polish version in its full form.[23]

Stage 1: Pilot implementation of the 2020 guidelines and their evaluation

The core element of the guidelines updating process was their pilot implementation and the assessment of their functioning and utility in a reference center. The general aim of this stage was to evaluate, in clinical practice, the functioning of the new Polish guidelines and their usefulness in assessing all key diagnostic dimensions in FASD: 1) PAE; 2) growth impairment; 3) dysmorphologies; 4) neurodevelopmental disorders.

Materials and methods

The study consisted of 2 parts. In the 1st part, the records of patients who presented for diagnostic purposes at the St. Louis Children's Hospital FASD Diagnostic Center in Kraków, Poland, between October 2020 and February 2022 were analyzed. The diagnosis was made by a team of professionals, including a physician (child psychiatrist and/or pediatrician) and a psychologist, according to the Polish criteria. In addition, a diagnosis was made according to the IOM criteria[11] based on the clinical data available at the time of analysis. During a 10-month period, 2 diagnostic teams used the guidelines: one with expertise in the field and another with expertise in childhood neurodevelopmental disorders, with minimal previous experience in the diagnosis of FASD. Before implementation, all team members participated in a structured 20-h training program delivered by the guideline co-authors (a physician and a psychologist), consisting of seminars and supervised practical exercises.

In the 2nd part of the study, qualitative data (interviews and written opinions) were collected from health professionals gathering information on PAE. These data provided insights into the practicality of the Polish guidelines and helped explain the differences in diagnosis based on IOM compared to Polish recommendations.

Sample

The records of 86 patients admitted to the clinic were analyzed, of whom 71 completed the diagnostic process. Fifteen patients did not complete the evaluation: some withdrew from the diagnostic process, some were lost to follow-up (the child's status changed and new caregivers refused to continue the evaluation), and others were referred to other facilities (genetic, neurometabolic, or autism spectrum disorder evaluation) for differential diagnosis, e.g., due to the biological mother's declaration that she had not consumed alcohol during pregnancy (confirmed no PAE, n = 4).

Among the patients, there were 40 boys and 31 girls. The median age was 96 months (range: 58–210 months). Ten children (14%) were raised by their biological parents (9 by their fathers and 1 by their mother), 31 (44%) were in adoptive families, 20 (28%) were in foster care, and 10 (14%) were living in institutional care (orphanage or childcare facility).

The 2nd group of study participants consisted of members of the FASD diagnostic team (n = 5) and professionals responsible for FASD screening ("screening experts"). The latter group (n = 8) comprised 2 family physicians, 1 neonatologist, 2 obstetrician–gynecologists, 2 school psychologists, and 1 educational psychologist, who applied the screening recommendations in their clinical or educational practice.

Measures

Data on PAE were collected using various methods: direct interview with the biological mother; interview with the biological mother's relatives or a representative of an institution (adoptive agency, social services, orphanage); information passed to the diagnostic team through other persons; and/or review of the child's health records. The evaluation of prenatal and postnatal growth impairment included all available data on birth weight as well as current and historical parameters (height and weight). The assessment of facial dysmorphic features was conducted by a clinical geneticist and pediatricians.

The neuropsychological assessment was adjusted to the child's age and needs. In the case of children up to 3 years of age, the Children's Development Scale (Dziecięca Skala Rozwojowa (DSR)) was used. For children aged 3 to 5 years, the Intelligence and Development Scales for Preschool Children, the Stanford–Binet Intelligence Scales (SB-5), and the Intelligence and Development Scales-2 (IDS-2) or Wechsler Intelligence Scale for Children (WISC–V) were used to assess emotional and social development.

For older children, the following tools were used: Intelligence and Development Scales (IDS-2), Stanford–Binet Intelligence Scales (SB-5), Raven's Progressive Matrices, Delis–Kaplan Executive Function System (D-KEFS), WISC–V, Benton Visual Retention Test (BVRT), Rey Auditory Verbal Learning Test (RVLT), and the D2 Test of Attention. If needed, additional clinical assessments were performed, and data from interviews and observations, psychiatric examinations, and results of assessments conducted by other psychologists (if available) were taken into account. All participating professionals (n = 13) provided feedback through written opinion essays and participated in a focus group interview moderated by a member of the project team.

Data analysis

All data were analyzed and discussed by the diagnostic team. In the first step, the team compared each patient's results in each diagnostic category with the Polish recommendations and classified the patient into one of the following categories: no FASD, FAS, ND-PAE, or FASD risk. The patient's results were then compared with the IOM diagnostic recommendations, and the patient was assigned to one of the following categories: FAS, pFAS, ARND, or no FASD.

Qualitative feedback from professionals was analyzed thematically to identify recurring facilitators and barriers to guideline implementation. Caregivers of children diagnosed during the implementation period were surveyed separately to obtain patient and family perspectives on the diagnostic process.

Results

In 14 (20%) cases, a diagnosis of FAS was made; 19 (27%) received a diagnosis of ND-PAE; 28 (39%) were considered "at risk of FASD"; and in 10 (14%) cases, FASD was excluded. According to the IOM criteria, 14 (20%) of the patients were diagnosed with FAS, 10 (15%) with pFAS, and 14 (21%) with ARND, while in 33 (42%) cases no FASD diagnosis was made (Table 1).

As shown in Table 1, the consistency of diagnoses based on the Polish[22] and IOM criteria[11] was 60%. The inconsistencies were due to differences in neurodevelopmental assessment and the inability to confirm PAE. The assessment of growth parameters and dysmorphic features yielded consistent diagnoses.

The guidelines were positively evaluated by professionals and perceived as clear, concise, and clinically useful. Participants particularly emphasized the value of standardized tools supporting both physical and psychological assessments. However, the neuropsychological criteria required to establish an FASD diagnosis were frequently described as relatively restrictive compared with other international guidelines, especially regarding the required number of affected brain domains. Based on the findings from the Stage 1 study, the need for further refinement of selected components of the guidelines was identified.

Stage 2: Structured literature review supporting the update

In December 2023, an evidence-informed literature review was conducted to support the update of the guidelines and to evaluate the current state of knowledge relevant to FASD diagnosis. The PubMed database was searched using the terms "FASD", "fetal alcohol spectrum disorders", and "prenatal alcohol exposure". A total of 150 results were retrieved. The search was further refined using targeted queries focused on dysmorphology and neuropsychological evaluation.

The identified publications were critically evaluated for clinical relevance and potential implications for physical and psychological assessment. Key findings and recent advances were summarized and presented to the working group to inform expert discussions and consensus-based revisions of the updated guidelines.

Stage 3: Formalized expert consensus process

The development of the updated guidelines was assigned to a multidisciplinary working group established on the initiative of the National Centre for Prevention of Addictions.

Table 1. Types of FASD diagnosis made according to the Polish (PL) and IOM's (US) recommendations

System/diagnosis	PL (n)	US (n)	Consistency	Causes of inconsistency
FAS	14	14	100%	–
ND-PAE	19	–	79%	Behavioral impairment without cognitive deficits Cognitive impairment only Impairment in <3 domains No functional impact in daily life PAE not confirmed
pFAS	–	10		
ARND	–	14		
FASD risk	28	–	–	
no diagnosis	10	33	30%	
All	71	71	60%	–

IOM – Institute of Medicine; FAS – fetal alcohol syndrome; ND-PAE – neurodevelopmental disorders associated with prenatal alcohol exposure; pFAS – partial fetal alcohol syndrome; ARND – alcohol-related neurodevelopmental disorders; FASD risk – fetal alcohol spectrum disorders risk; PAE – prenatal alcohol exposure.

The composition of the working group reflected the interdisciplinary nature of FASD diagnosis and included 6 physicians (a geneticist, pediatric neurologist, pediatric psychiatrist, pediatrician, pediatric gastroenterologist, and obstetrician–gynecologist), 4 psychologists, 1 pedagogue/social therapist, and 2 caregiver representatives.

Due to the limited number of national experts with extensive clinical experience in FASD diagnosis in Poland, the consensus process was based on a repeated expert panel methodology, consistent with an expert opinion consensus approach. The panel worked in an iterative manner, with successive rounds of discussion and refinement of the proposed amendments.

Draft amendments to specific sections of the guidelines were prepared by domain experts according to their areas of clinical expertise (pediatricians and geneticists – growth assessment and evaluation of dysmorphic features; child and adolescent psychiatrist, pediatric neurologist, and neuropsychologist – central nervous system (CNS) evaluation; obstetrician – PAE assessment; psychologist – diagnostic scheme; psychologist – diagnostic categories). These proposals were subsequently discussed during structured virtual and in-person meetings of the working group. An additional element of the guideline development process was a review by a psychologist, the first author of the previous edition of the guidelines.

Consensus was achieved through discussion, incorporation of feedback from all panel members, and revision of recommendations until agreement was reached. In cases of initial disagreement, alternative formulations were discussed and refined in subsequent rounds until a shared position was established.

Stage 4: Stakeholder consultation

A stakeholder meeting was held. The proposed guidelines were presented to the representatives of the diagnostic centers in Poland, a broader discussion was initiated, and the questions were answered.

Results of the process of the guidelines updating

The updated guidelines include the following sections: Introduction, Methodology, Diagnostic categories, Diagnostic scheme, Ethical considerations, Evaluation of PAE, Evaluation of facial dysmorphism, Evaluation of growth impairment, and Evaluation of the central nervous system.

Compared to the 2020 guidelines, a section describing Ethical considerations was added. The list of annexes was reviewed, and 7 annexes were included in the final document.

Annex 1: Feeding problems in infants and toddlers that require evaluation by a speech therapist.

Annex 2: Screening for CNS impairment suggesting the need for referral to a pediatric neurologist.

Annex 3: Behavioral and cognitive characteristics suggesting the need for referral for FASD diagnosis.

Annex 4: Differential diagnosis of FASD.

Annex 5: Structured anamnesis with the birth mother.

Annex 6: A list of standardized tools for neuropsychological evaluation.

Annex 7: Psychiatric evaluation in the diagnosis of FASD.

In all chapters, new information from the literature review was added to illustrate the current state of the art, presenting the crucial amendments made compared to the 2020 guidelines. The full printed guidelines are available.[23]

Diagnostic categories

In the 2020 guidelines, 3 diagnostic categories were recommended: FAS, ND-PAE, and a descriptive category "at risk of FASD." The aforementioned categories were maintained in the new guidelines (Table 1). However, a new category was added: partial fetal alcohol syndrome (pFAS) (Table 1), which can be diagnosed in cases of both confirmed and non-confirmed PAE (Table 2).

Table 2. Categories of FASD proposed in the updated 2025 Polish guideline

Category	FASD			At risk of FASD
	FAS	pFAS	ND-PAE	
Prenatal alcohol exposure	yes or unknown	yes or unknown	yes	Confirmed PAE, or unknown if all three sentinel facial features are present
Pre- and/or postnatal growth deficits	yes	yes/no yes	N/A	to be observed
Sentinel facial features: Short palpebral fissures Thin upper lip vermilion border Smooth (flat) philtrum	3 of 3	2 of 3	N/A	to be observed
Neurodevelopmental disorders	One key neurological symptom; Deficits in ≥3 cognitive domains; ≥1 symptom in each domain (cognitive, self-regulation, adaptive functioning).			to be observed

FASD – fetal alcohol spectrum disorders; FAS – fetal alcohol syndrome; ND-PAE – neurodevelopmental disorders associated with prenatal alcohol exposure; pFAS – partial fetal alcohol syndrome; ARND – alcohol related neurodevelopmental disorders; N/A – not applicable.

Diagnostic scheme

Consequently, the diagnostic scheme (Fig. 1) was reviewed and updated to include the diagnosis of pFAS. The term "functional diagnosis" was removed from the scheme, as it was considered an integral part of the diagnosis itself; however, it was clarified that functional assessment is not equivalent to the diagnosis of FASD. The role of standardized tools was emphasized not only during the diagnostic process but also during screening. The pFAS category was retained for pragmatic reasons. In the Polish clinical and legal context, it allows for the classification of children with clear dysmorphic features who do not meet the full criteria for FAS but still require support.

This ensures that such children are not left without a diagnosis or access to adequate services. The authors emphasize that this is a transitional solution that can be revised in the future with broader implementation of ICD-11 and ND-PAE. However, the "at risk of FASD" category was maintained because the authors agreed that some children may be missing components of key diagnostic criteria at the time of presentation in the diagnostic setting. For these patients, it was considered beneficial to monitor them in case they meet the diagnostic criteria in the future.

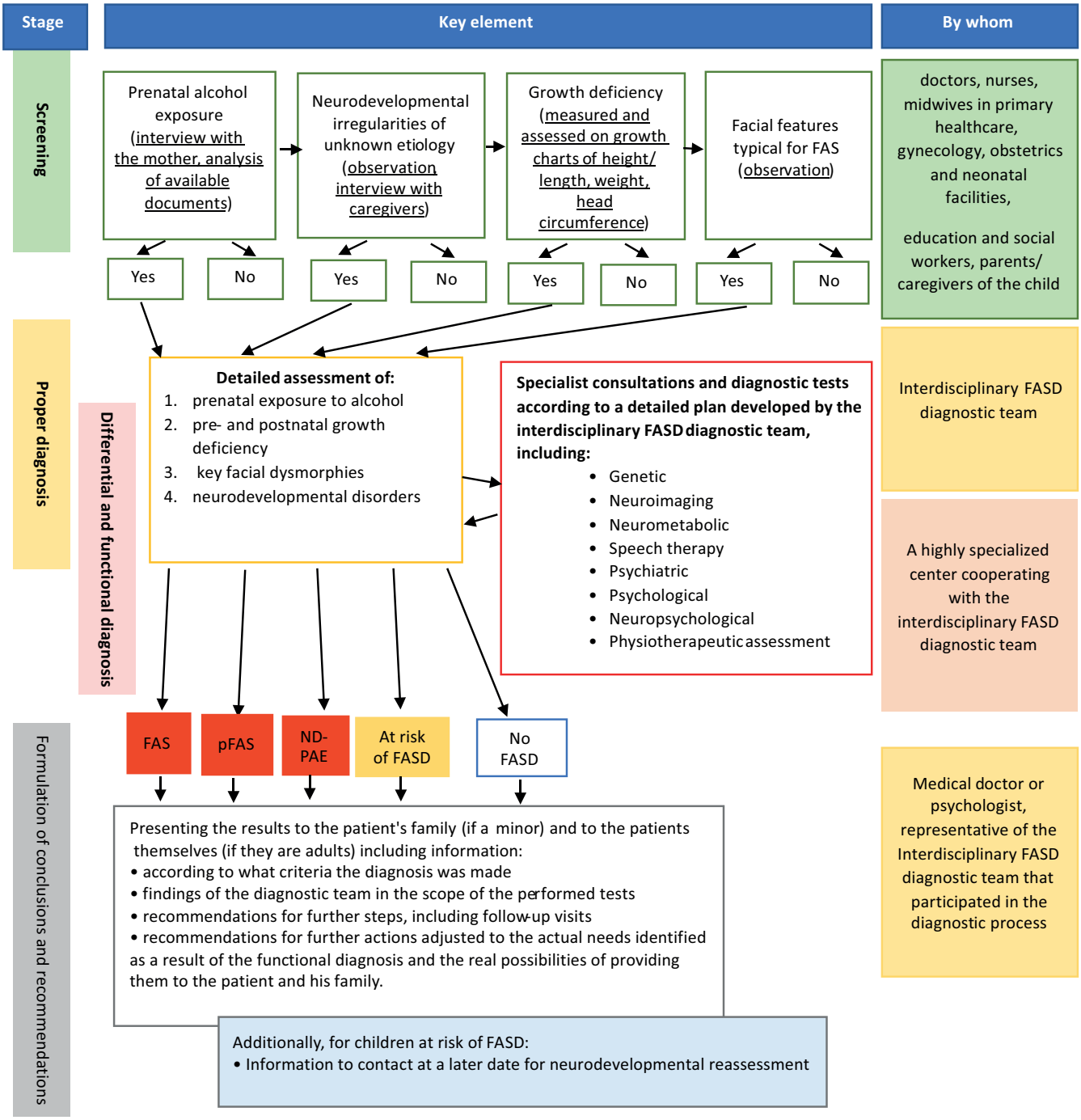


Fig. 1. Diagnostic scheme

Evaluation of prenatal alcohol exposure

According to the 2020 guidelines, there were 3 methods to evaluate PAE[22]: direct anamnesis obtained from the birth mother; indirect anamnesis obtained from a person who accompanied the birth mother during her pregnancy; and medical, judicial, or employment records, as well as broadly understood social welfare documentation. Due to the difficulty of obtaining reliable answers to questions about alcohol consumption during pregnancy through the interview method,[5] a new section was added in the 2025 guidelines to the chapter illustrating the role of biomarkers in the evaluation of PAE.

Ethyl glucuronide (EtG) and phosphatidylethanol (PEth) were considered the 2 biomarkers with the most established role in confirming PAE. However, PEth can be detected in maternal blood only within a limited time frame of 4–6 weeks after alcohol consumption.[24] On the other hand, EtG accumulates in human hair, and its detection can reflect alcohol consumption throughout pregnancy across all 3 trimesters.[25] The promising role of biomarkers in the evaluation of PAE was highlighted; however, information regarding the requirement for mandatory maternal consent was also included.

Evaluation of facial dysmorphism

The evaluation of facial dysmorphism in the 2020 guidelines involved palpebral fissure measurement (using 2 methods: traditional with a ruler and a computerized method) and lip/philtrum assessment with the lip/philtrum guide.[22] The aforementioned procedure was maintained in the 2025 guidelines. However, in the new guidelines, the use of only one selected method for assessing the length of the palpebral fissure was approved. In addition, a table with more subtle dysmorphic features was added to illustrate the complexity of the physical examination in the FASD diagnostic process.

Evaluation of growth impairment

The evaluation of prenatal and postnatal growth impairment was included in the 2020 guidelines. It involved birthweight and current and historical parameters (height and weight) analysis according to growth charts. The procedure remained unchanged in the 2025 guidelines. However, an amendment regarding growth charts was made: after the age of 3, the use of the 2010 Polish (OLA/OLAF) growth charts[26] was recommended instead of the 1999 growth charts previously recommended.[27]

Evaluation of central nervous system

The evaluation of the CNS was the most comprehensive revision of the chapter. The levels of evaluation (neurological, neurodevelopmental, and neurobehavioral) were specified. Each level was characterized by potential pathologies that may indicate CNS impairment. In the neurological evaluation, microcephaly, CNS malformations observed in previously performed imaging, or epilepsy of prenatal origin were considered key neurological symptoms. However, it was underlined that the presence of features of cerebral palsy, focal symptoms, central deafness, or other neurological symptoms indicates the need for consultation with a pediatric neurosurgeon.

Neurodevelopmental evaluation encompasses the assessment of motor, sensory, and perceptual functions, as well as the assessment of the articulatory mechanism and oral praxis, primary orofacial functions, respiratory pattern, and the quality of speech and communicative competencies. Neurobehavioral evaluation consists of the assessment of the following domains: cognition (intelligence, executive functions, learning, memory, visuospatial competencies), self-regulation (emotional regulation, attention, impulse control), and adaptive functioning (communication, social cognition, motor functions, everyday coping strategies).

The evaluation of the 3 levels should be accompanied by a functional analysis, including the assessment of psychopathological symptoms and functioning in home and school settings. A section on concomitant psychopathological diagnoses was added. Significant CNS impairment was defined as the presence of: a) 1 key neurological symptom (microcephaly, CNS malformation, or epilepsy of prenatal origin); b) deficits in ≥3 neurocognitive areas; or c) ≥1 symptom in each domain (neurocognitive, self-regulation, adaptive functioning).

Discussion

The Polish FASD diagnostic criteria developed in 2020, although generally positively received, were poorly known among specialists and criticized mainly for the requirements for diagnosing PAE, which were difficult to apply in clinical practice.[24] Pilot implementation of these guidelines (stage 1 of the updating procedure) largely confirmed both the expectations and concerns of Polish experts related to the assessment of PAE. First of all, in retrospective studies, i.e., when the diagnostic procedure is conducted in the case of a child aged several or more than a dozen years, collecting reliable information about the alcohol consumption of the pregnant mother is very difficult and often impossible. As a result, if PAE is considered "probable", patients with neurodevelopmental disorders but without characteristic facial features are, according to the 2020 Polish recommendations, classified as "at risk" of FASD, whereas according to the IOM criteria, FASD is ruled out in these patients.

Another cause of inconsistent diagnoses, depending on the recommendations used, was the criteria for assessing the CNS. The 2020 Polish criteria were more restrictive in this regard than the IOM criteria. As a result, diagnostic consistency between the recommendations was

60%, which can be considered high. In studies comparing diagnoses based on the IOM criteria and the 4-Digit Code, consistency was 38%.[28]

In the absence of a single diagnostic or imaging test to confirm FASD, diagnostic criteria remain the gold standard. The proposed 2025 Polish guidelines are consistent with international recommendations; however, they also represent an attempt to adapt existing scientific data to the Polish context (e.g., specific growth charts, validated psychometric tools for evaluation). From a content perspective, the Polish recommendations are in agreement with other widely used FASD diagnostic guidelines. Regarding diagnostic categories, they largely reproduce the nomenclature from the guidelines developed by Hoyme et al.[11]

However, alcohol-related neurodevelopmental disorder (ARND) is instead termed ND-PAE to avoid a definitive causal interpretation. Nevertheless, this category does not fully overlap with DSM-5 ND-PAE,[29,30] in which a diagnosis of ND-PAE can be made with or without the presence of dysmorphic features. More similarities can be found between the definition in the 2025 Polish guidelines and the Canadian guidelines.[14] The addition of a new diagnostic category, pFAS, may be viewed as controversial. However, the authors return to a system in which children presenting with 2 out of 3 dysmorphic features are treated as a separate group.

The category was added following suggestions from professionals who pointed out that patients presenting with 2 dysmorphic features, growth restriction, and neurobehavioral problems but with unknown PAE are left without a formal diagnosis. The "at risk" category was maintained, as professionals underlined its utility in cases where some information is lacking. Children assigned to this category remain in the diagnostic process and can be transferred to other categories if new information emerges, such as confirmed PAE (previously unknown) or the development of neurobehavioral problems.

The dysmorphological evaluation and its principles do not differ substantially among the available diagnostic guidelines, which are based on the lip and philtrum guide and palpebral fissure measurements.[11,12,14] An important difference between the 2025 and 2020 Polish guidelines is that the process was simplified by allowing only 1 evaluation method instead of two. However, no other diagnostic guidelines[11,12,14] provide for double verification of palpebral fissure length, and, in the opinion of professionals, this approach hindered the implementation of the guidelines.

Neuropsychological evaluation is a key component of FASD diagnosis. The undisputed advantage of the Polish guidelines is that they are accompanied by a list of validated and available psychometric tools recommended in Poland. In the amended recommendations, the role of structural brain defects such as microcephaly is emphasized. This is especially important in diagnosing young children, among whom subtle neuropsychological symptoms may not be evident. Regarding alcohol exposure, the 2025 Polish guidelines, in light of the latest research, are the first to suggest the use of biomarkers in the process of PAE assessment.

Furthermore, the path to population screening is still distant, and biomarkers need to be evaluated in terms of cost-effectiveness and technological availability. Therefore, the use of biomarkers remains a methodologically justified possibility that may be incorporated into the diagnostic process following further expansion of medical technology. Moreover, the use of biomarkers is only possible with the informed consent of the mother, which is consistent with previous international publications on this matter.[8,31]

The authors of the 2025 guidelines emphasize that they should be reviewed and updated within a 5-year period; however, their applicability will be continuously monitored. The 2020 guidelines were created in the hope that they would be the first step toward establishing an integrated system of care for children with FASD in Poland. The authors of the 2025 guidelines emphasize that systemic, government-coordinated actions are urgently needed, as relatively few steps have been taken in the past 5 years and FASD diagnosis is still not covered by the public healthcare system. Although data on alcohol consumption among Polish women are alarming,[4] and only some preventive efforts seem to be effective,[6] FASD remains a marginalized health issue. The authors of the guidelines believe that the updated version will be widely distributed not only to diagnostic centers but also to primary care offices, regional centers, and obstetric/neonatology departments. Within the framework of the project funded by the National Centre for Prevention of Addictions, printed copies of the guidelines were distributed nationwide.

In the 2025 guidelines, an effort has been made to simplify the diagnostic pathway – this could potentially reduce the entry barrier for new diagnostic settings. An additional advantage of the 2025 guidelines is their potential to improve daily clinical practice. By introducing a simplified diagnostic process and recommending standardized tools available in Poland, they can streamline diagnostic work not only in referral centers but also in smaller facilities. This, in turn, could reduce regional disparities in access to FASD diagnosis in Poland (Table 3).

Limitations

Although the update of the 2020 guidelines[22] is an important step toward improving the system for addressing FASD-related problems, the limitations of our work must be acknowledged. As recommended in the 2020 guidelines, evaluation in the clinical setting was the first step in verifying the guidelines (Stage 1). However, feedback was collected at only a single specialized facility. Therefore, it is uncertain to what extent the conclusions from this study can be generalized.

Table 3. Changes in the Polish FASD guidelines (2025 update) compared with the 2020 guidelines

Category	2020 guidelines	2025 guidelines
Diagnostic categories	• FAS • ND-PAE • FASD risk	FAS, ND-PAE, FASD risk maintained, pFAS added
Evaluation of prenatal alcohol exposure	Three methods of PAE determination (direct and indirect history taking, and documentation review).	Three methods of PAE determination maintained; information on the potential use of PAE biomarkers added.
Evaluation of facial dysmorphism	• Two measurements with each of 2 methods (manual with a ruler and with a computer) • Lip assessment according to the Lip-Philtrum Guide • Philtrum assessment according to the Lip-Philtrum Guide	• Two measurements with 1 selected method (manual with a ruler and with a computer) • Lip evaluation according to the Lip-Philtrum Guide Philtrum assessment according to the Lip-Philtrum Guide
Evaluation of growth impairment	Analysis of birthweight, height and weight at the time of diagnosis according to: • WHO growth charts for term children aged 0–5 years and Fenton growth charts for preterm infants • Palczewska and Niedźwiedzka growth charts >5[28]	Analysis of birthweight, height, and weight at the moment of diagnosis according to: • WHO growth charts for term children aged 0–3 years and Fenton growth charts for preterm infants; • OLA/OLAF growth charts[26] >3 years.
Evaluation of the central nervous system (CNS)	CNS impairment is defined as: Deficits in ≥3 cognitive domains (≥2 if neurological symptoms are present); Abnormalities in ≥3 domains of emotional, social, or adaptive functioning, or psychopathological symptoms; Clinically significant impairment in daily functioning and school performance (or occupational functioning in adults).	CNS impairment defined as: 1 key neurological symptom; Deficits in ≥3 cognitive domains; ≥1 symptom in each domain (cognitive, self-regulation, adaptive functioning).

FASD – fetal alcohol spectrum disorders; FAS – fetal alcohol syndrome; ND-PAE – neurodevelopmental disorders associated with prenatal alcohol exposure; pFAS – partial fetal alcohol syndrome.

Second, although a literature search was conducted to update the guidelines with the most recent evidence and methods (Stage 2), most modifications were made based on feedback from professionals in the field (Stage 3). Therefore, in the summary of the 2025 guidelines, it was suggested that future updates should be entrusted to scientific societies.

Conclusions

The 2025 guidelines have been positively received as a unified diagnostic tool, and their implementation in clinical practice has provided valuable feedback that prompted the re-evaluation of several recommendations. Based on this experience, the updated guidelines are now expected to be widely disseminated and supported by structured training for health professionals, which may further reduce the diagnostic gap, improve consistency between centers, and strengthen interdisciplinary collaboration, thereby increasing momentum for systemic solutions to this complex health issue.

The 2025 Polish guidelines not only standardize diagnostic procedures but also set the stage for coordinated national strategies in the prevention, diagnosis, and care of individuals affected by FASD.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo at https://doi.org/10.5281/zenodo.18470451.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Katarzyna A. Dyląg https://orcid.org/0000-0001-6886-0136
Małgorzata Klecka https://orcid.org/0009-0002-7000-6507
Iwona Palicka https://orcid.org/0000-0002-8779-1946
Katarzyna Okulicz-Kozaryn https://orcid.org/0000-0002-7981-3885
Agata Cichoń-Chojnacka https://orcid.org/0000-0002-0451-8689
Tomasz Maciejewski https://orcid.org/0000-0003-3761-1924
Robert S. Śmigiel https://orcid.org/0000-0003-2930-9549

References

1. Lange S, Probst C, Gmel G, Rehm J, Burd L, Popova S. Global prevalence of fetal alcohol spectrum disorder among children and youth: A systematic review and meta-analysis. *JAMA Pediatr.* 2017;171(10):948. doi:10.1001/jamapediatrics.2017.1919
2. Roozen S, Peters GY, Kok G, Townend D, Nijhuis J, Curfs L. Worldwide prevalence of fetal alcohol spectrum disorders: A systematic literature review including meta-analysis. *Alcohol Clin Exp Res*. 2016; 40(1):18–32. doi:10.1111/acer.12939
3. Okulicz-Kozaryn K, Borkowska M, Brzózka K. FASD prevalence among schoolchildren in Poland. *J Appl Res Intellect Disabil*. 2017;30(1):61–70. doi:10.1111/jar.12219
4. Okulicz-Kozaryn K, Terlikowska J, Brzózka K, Borkowska M. Prevention and intervention for FASD in Poland. *J Pediatr Neuropsychol*. 2017; 3(1):79–92. doi:10.1007/s40817-016-0025-9
5. Okulicz-Kozaryn K, Marchei E, Helwich E, et al. The prevalence and changes in alcohol consumption across three trimesters of pregnancy assessed by ethyl glucuronide concentration in maternal hair and self-reports: A cross-sectional study. *Eur Addict Res*. 2024;30(6):378–389. doi:10.1159/000542474

6. Okulicz-Kozaryn K. Is public health response to the phenomenon of alcohol use during pregnancy adequate to the Polish women's needs? *Int J Environ Res Public Health*. 2022;19(8):4552. doi:10.3390/ijerph19084552
7. Popova S, Dozet D, Shield K, Rehm J, Burd L. Alcohol's impact on the fetus. *Nutrients*. 2021;13(10):3452. doi:10.3390/nu13103452
8. Kable JA, Jones KL. Identifying prenatal alcohol exposure and children affected by it: A review of biomarkers and screening tools. *Alcohol Res*. 2023;43(1):3. doi:10.35946/arcr.v43.1.03
9. Chabenne A, Moon C, Ojo C, Khogali A, Nepal B, Sharma S. Biomarkers in fetal alcohol syndrome. *Biomarkers Genom Med*. 2014;6(1):12–22. doi:10.1016/j.bgm.2014.01.002
10. Jones KL, Smith DW. Recognition of the fetal alcohol syndrome in early infancy. *Lancet*. 1973;302(7836):999–1001. doi:10.1016/s0140-6736(73)91092-1
11. Hoyme HE, Kalberg WO, Elliott AJ, et al. Updated Clinical Guidelines for Diagnosing Fetal Alcohol Spectrum Disorders. *Pediatrics*. 2016;138(2):e20154256. doi:10.1542/peds.2015-4256
12. Astley SJ. Validation of the fetal alcohol spectrum disorder (FASD) 4-Digit Diagnostic Code. *J Popul Ther Clin Pharmacol*. 2013;20(3):e416–467. PMID:24323701.
13. Chudley AE, Conry J, Cook JL, et al. Fetal alcohol spectrum disorder: Canadian guidelines for diagnosis. *Can Med Assoc J*. 2005;172(5 Suppl):S1–S21. doi:10.1503/cmaj.1040302
14. Cook JL, Green CR, Lilley CM, et al. Fetal alcohol spectrum disorder: A guideline for diagnosis across the lifespan. *Can Assoc Radiol J*. 2016;188(3):191–197. doi:10.1503/cmaj.141593
15. Coles CD, Gailey AR, Mulle JG, Kable JA, Lynch ME, Jones KL. A comparison among 5 methods for the clinical diagnosis of fetal alcohol spectrum disorders. *Alcohol Clin Exp Res*. 2016;40(5):1000–1009. doi:10.1111/acer.13032
16. Brown JM, Bland R, Jonsson E, Greenshaw AJ. The standardization of diagnostic criteria for fetal alcohol spectrum disorder (FASD): Implications for research, clinical practice and population health. *Can J Psychiatry*. 2019;64(3):169–176. doi:10.1177/0706743718777398
17. Landgraf M, Schmucker C, Heinen F, Ziegler A, Kopp I, Strieker S. Diagnosis of fetal alcohol spectrum disorders: German guideline version 2024. *Eur J Paediatr Neurol*. 2024;53:155–165. doi:10.1016/j.ejpn.2024.11.002
18. Landgraf MN, Nothacker M, Heinen F. Diagnosis of fetal alcohol syndrome (FAS): German guideline version 2013. *Eur J Paediatr Neurol*. 2013;17(5):437–446. doi:10.1016/j.ejpn.2013.03.008
19. Regional Competence Center for children with prenatal alcohol and/or drug exposure, South-Eastern Health Region (RK-MR HSØ). Norwegian Clinical Guideline for Diagnostic Assessment of Fetal Alcohol Spectrum Disorder (FASD) in Childhood and Adolescence. Andal, Norway: Sørlandet Hospital; 2024. https://www.sshf.no/4a77d3/contentassets/7b4559a3bcbb45c7a5bed5d713116c37/norwegian-fasd-guideline-2024-final.pdf.
20. Micangeli G, Menghi M, Paparella R, et al. Italian guidelines for the diagnosis and treatment of fetal alcohol spectrum disorders: Diagnostic criteria. *Riv Psichiatr*. 2024;59(5):195–202. doi:10.1708/4360.43509
21. Watkins RE, Elliott EJ, Wilkins A, et al. Recommendations from a consensus development workshop on the diagnosis of fetal alcohol spectrum disorders in Australia. *BMC Pediatr*. 2013;13(1):156. doi:10.1186/1471-2431-13-156
22. Okulicz-Kozaryn K, Maryniak A, Borkowska M, Śmigiel R, Dylag KA. Diagnosis of Fetal Alcohol Spectrum Disorders (FASDs): Guidelines of Interdisciplinary Group of Polish Professionals. *Int J Environ Res Public Health*. 2021;18(14):7526. doi:10.3390/ijerph18147526
23. Borkowska M, Cichoń-Chojnacka A, Dyląg KA, et al. Rozpoznawanie spektrum płodowych zaburzeń alkoholowych: Wytyczne opracowane przez interdyscyplinarny zespół polskich ekspertów. *Medycyna Praktyczna – Pediatria*. 2025;1(wydanie specjalne):1–52. https://adst.mp.pl/s/www/pediatria/WS-Pediatria-2025-01_FASD-24-09.pdf.
24. Finanger T, Spigset O, Gråwe RW, et al. Phosphatidylethanol as blood biomarker of alcohol consumption in early pregnancy: An observational study in 4,067 pregnant women. *Alcohol Clin Exp Res*. 2021;45(4):886–892. doi:10.1111/acer.14577
25. Montag A. Fetal alcohol-spectrum disorders: Identifying at-risk mothers. *Int J Womens Health*. 2016;8:311–323. doi:10.2147/IJWH.S85403
26. Kułaga Z, Różdżyńska A, Palczewska I, et al. Percentile charts of height, body mass and body mass index in children and adolescents in Poland: Results of the OLAF study. *Medycyna Praktyczna – Pediatria*. 2010;7:690–700. https://www.standardy.pl/artykuly/id/122%20file.
27. Palczewska I, Niedzwiedzka Z. Somatic development indices in children and youth of Warsaw [in Polish]. *Med Wieku Rozwoj*. 2001;5(2 Suppl 1):18–118. PMID:11675534.
28. Astley SJ, Bledsoe JM, Davies JK, Thorne JC. Comparison of the FASD 4-Digit Code and Hoyme et al. 2016 FASD diagnostic guidelines. *Adv Pediatr Res*. 2017;4(3):13. doi:10.12715/apr.2017.4.13
29. Kable JA, O'Connor MJ, Olson HC, et al. Neurobehavioral disorder associated with prenatal alcohol exposure (ND-PAE): Proposed DSM-5 diagnosis. *Child Psychiatry Hum Dev*. 2016;47(2):335–346. doi:10.1007/s10578-015-0566-7
30. American Psychiatric Association (APA). *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, D.C, USA: American Psychiatric Association (APA); 2013. doi:10.1176/appi.books.9780890425596
31. Zizzo N, Di Pietro N, Green C, Reynolds J, Bell E, Racine E. Comments and reflections on ethics in screening for biomarkers of prenatal alcohol exposure. *Alcohol Clin Exp Res*. 2013;37(9):1451–1455. doi:10.1111/acer.12115

Global, regional, and national burden of gastroesophageal reflux disease (GERD), 1990–2021, with projections to 2050

Kejin Li[1,C,D], Hui Li[1,E,F], Zhentao An[1,B,C], Xiangxiang Xu[1,B,C], Jing Zuo[1,B,C], Jiaxin Li[1,C], Xueli Qian[1,C], Liu Liu[2,B,C], Jingjing Cui[3,C]

[1] Department of Gastroenterology, Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, China
[2] Endocrine and Diabetes Center, Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, China
[3] Department of Oncology, Affiliated Hospital of Integrated Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1073–1084

Address for correspondence
Kejin Li
E-mail: lkj18725477619@163.com

Funding sources
This study was supported by the National Construction of Key Specialized Clinics in Traditional Chinese Medicine (grant No. 20240110201).

Conflict of interest
None declared

Received on May 15, 2025
Reviewed on August 5, 2025
Accepted on September 5, 2025

Published online on June 23, 2026

Abstract

Gastroesophageal reflux disease (GERD) is a common digestive system disorder encountered in clinical practice. In recent years, the prevalence of GERD has increased, substantially affecting patients' daily lives. This study aimed to address this issue by leveraging data from the Global Burden of Disease (GBD) 2021 database to systematically evaluate the burden and epidemiological characteristics of GERD and to predict trends in GERD burden from 2022 to 2050. We comprehensively analyzed the burden of GERD from 1990 to 2021, evaluated years lived with disability (YLDs), prevalence and incidence rates, and conducted stratified analyses according to geographical region, Sociodemographic Index (SDI), sex, and age groups ranging from 0 to 95 years. Globally, in 2021, there were 324 million (95% uncertainty interval (95% UI): 287.7–358.9 million) incident cases, 826 million (95% UI: 733–926 million) prevalent cases, and 6.34 million (95% UI: 3.19–11.24 million) YLDs attributable to GERD. Among individuals aged 0–95 years, women aged 35–39 years had the highest prevalence, estimated at 45.32 million (95% UI: 32.92–61.05 million). The burden was highest in middle-SDI regions and lowest in high-SDI regions. Predictions using autoregressive integrated moving average (ARIMA) modeling indicated that the global burden of GERD will continue to increase from 2022 to 2050, posing increasingly severe challenges to global healthcare systems, particularly among women. It is projected that by 2050, women will account for 54% of new cases.

Key words: epidemiology, gastroesophageal reflux, Global Burden of Disease

Cite as
Li K, Li H, An Z, et al. Global, regional, and national burden of gastroesophageal reflux disease (GERD), 1990–2021, with projections to 2050. *Adv Clin Exp Med*. 2026;35(6):1073–1084. doi:10.17219/acem/210367

DOI
10.17219/acem/210367

Highlights

- In 2021, gastroesophageal reflux disease (GERD) affected 826 million people globally, with 324 million new cases and 6.34 million years lived with disability (YLDs), according to Global Burden of Disease (GBD) 2021 data.
- Women aged 35–39 years showed the highest prevalence, with 45.32 million cases, and the burden was greatest in regions with a medium Socio-Demographic Index (SDI).
- Autoregressive Integrated Moving Average (ARIMA) modeling predicts a continued increase in the burden of GERD through 2050, especially among women, who are projected to account for 54% of new cases.
- These findings highlight the growing public health impact of GERD and the need for targeted prevention strategies across age, sex, and SDI groups.

Introduction

Gastroesophageal reflux disease (GERD) is a prevalent gastrointestinal disorder worldwide, pathologically defined by the retrograde flow of gastric contents into the esophagus, oropharynx, or pulmonary system. Clinical manifestations include typical symptoms (heartburn and acid regurgitation) as well as atypical extraesophageal presentations, including laryngitis, chronic cough, and asthma exacerbations. The disease exhibits a bimodal age distribution, affecting both pediatric and adult populations. In children aged <8 years, GERD primarily manifests as feeding refusal, recurrent vomiting, and pulmonary complications (chronic cough and recurrent pneumonia), whereas older children (>8 years) and adolescents typically present with adult-type symptoms such as epigastric pain, nocturnal cough, and wheezing.[1,2]

Endoscopic classification distinguishes 2 phenotypes: erosive esophagitis (EE; 30% prevalence), characterized by mucosal breaks (Los Angeles Grade A–D), and nonerosive reflux disease (NERD; 60–70% prevalence), characterized by intact mucosa despite symptomatic reflux.[3] Barrett's esophagus (BE) is a specific form of GERD characterized by replacement of the squamous epithelium in the lower esophagus with metaplastic columnar epithelium. Patients with BE may experience reflux symptoms, and some may also present with dysphagia and retrosternal pain, although a significant proportion remains asymptomatic. Barrett's esophagus is one of the major complications of GERD and the only recognized precancerous lesion of esophageal adenocarcinoma (EAC).[4,5]

The epidemiological burden of GERD is greater among individuals aged ≥50 years, smokers, users of nonsteroidal anti-inflammatory drugs (NSAIDs), and patients with obesity.[6,7] Research has shown that the diagnosis, treatment, and cancer surveillance associated with GERD impose considerable economic and psychological burdens. Currently, several studies have analyzed the epidemiological burden of GERD in different regions based on the Global Burden of Disease (GBD) 2017 and GBD 2019 databases. However, these studies often rely on relatively outdated data and still lack analyses stratified by sex and age groups, as well as projections of disease trends over the next 30 years.[8,9] Given the dynamic changes in the burden of GERD over time, it is important to promptly understand the current epidemiological status of GERD and monitor its changing trends. This knowledge is crucial for developing effective intervention strategies and policies at the global, regional, and national levels.

To provide comparable and up-to-date information on the burden of GERD, this study used the latest GBD 2021 data to analyze the incidence, prevalence, and years lived with disability (YLDs) of GERD across different regions, sexes, and age groups worldwide. In addition, the incidence of GERD from 2022 to 2050 was projected to provide a reference for public health interventions and policy development in different countries and regions and to support further basic research and clinical practice related to GERD.

Materials and methods

Data source

The GBD 2021 database expands epidemiological surveillance of GERD across 811 subnational regions in 204 countries. All data resources were obtained from the GBD Results Tool (http://ghdx.healthdata.org/gbd-results-tool) of the Global Health Data Exchange (GHDX) platform, maintained by the Institute for Health Metrics and Evaluation (IHME; Washington, D.C., USA). The diagnosis of GERD was classified according to the 10th revision of the International Classification of Diseases (ICD-10) using codes R12.11, K21–K21.9, and K22.7–K22.719.[10] Age-standardized rate estimates and counts per 100,000 population are presented according to the GBD standard population structure.

Within the GBD framework, 95% uncertainty intervals (95% UIs) are provided for all estimates. Final estimates were calculated as the mean of 1,000 estimates obtained through sampling, with the lower and upper 95% UIs corresponding to the 2.5th and 97.5th percentiles, respectively, among the 1,000 samples.[11] Confidence intervals (CIs)

were calculated based on standard errors (SEs) under the assumption of a normal distribution. Confidence intervals are used to quantify uncertainty surrounding the estimation of a single model parameter and are commonly applied to assess the statistical significance of trends. Accordingly, in GBD studies, UIs are used to characterize the magnitude of disease burden estimates, whereas CIs are applied to evaluate changes in disease burden trends over time.

Additionally, the Sociodemographic Index (SDI), a composite measure of income, education, and fertility reflecting the level of sociodemographic development in a country or territory, was used. The SDI incorporates 5 levels based on SDI quintiles: low, low-middle, middle, high-middle, and high. The GBD data are available for 20 non-overlapping 5-year age groups, ranging from under 5 years to over 95 years.

Trend analysis

We explored the temporal trends in the global burden of GERD from 1990 to 2021 according to age group, sex, SDI, GBD region, and country, and estimated the estimated annual percentage change (EAPC) using a linear regression model. Based on the EAPC values, we assessed disease burden trends in each GBD region and identified regions with similar patterns of disease burden change.[12]

Predictive analysis

To better inform public health policy and healthcare resource allocation, we further projected the global burden of GERD through 2050. Autoregressive Integrated Moving Average (ARIMA) models are statistical models commonly used for time-series analysis and forecasting.[13] These models capture the characteristics of time-series data and predict disease trends over the next 30 years by combining autoregressive (AR), differencing (I), and moving average (MA) components.

Results

Trends in the incidence of GERD in 2021

In 2021, the global number of incident cases reached 324 million (95% UI: 288–359 million), reflecting an 80% increase from 1990. After adjustment for age structure, the global age-standardized incidence rate (ASIR) increased from 3,740 per 100,000 population (95% UI: 3,314–4,142) in 1990 to 3,882 per 100,000 (95% UI: 3,446–4,300) in 2021, with an EAPC of 0.36 (95% CI: 0.29–0.43).

The SDI-stratified analysis revealed declining trends in high-SDI (EAPC = −0.13; 95% CI: −0.20 to −0.06), high-middle SDI (EAPC = −0.23; 95% CI: −0.30 to −0.15), and low-SDI regions (EAPC = −0.01; 95% CI: −0.01 to −0.01), whereas middle-SDI regions exhibited an increasing trend (EAPC = 0.28; 95% CI: 0.26–0.31).

Regionally, ASIRs varied widely, ranging from 6,248 per 100,000 (95% UI: 5,568–6,875) in Tropical Latin America to 1,850 per 100,000 (95% UI: 1,610–2,090) in East Asia, with the highest incidence rates (>6,000 per 100,000) observed in Latin America (Tropical Latin America, Central Latin America, and Andean Latin America) and the Caribbean. In contrast, the lowest rates (<3,000 per 100,000) were found in East Asia, the Western Pacific region, Oceania, Southeast Asia, and High-Income Asia Pacific. Notably, Asia (EAPC = 0.60; 95% CI: 0.50–0.70) and High-Income Asia Pacific (EAPC = 0.50; 95% CI: 0.15–0.86) experienced pronounced growth, contrasting with the decline observed in High-Income North America (EAPC = −0.30; 95% CI: −0.55 to −0.06).

At the national level, the highest burdens were observed in India (76,322,880; 95% UI: 67,266,798–84,678,866), China (32,387,866; 95% UI: 27,851,900–36,711,150), and the USA (16,021,759; 95% UI: 14,111,407–17,823,288). The 3 countries with the lowest ASIRs in 2021 were Norway (1,838.34; 95% UI: 1,582.47–2,072.53), China (1,844.31; 95% UI: 1,605.06–2,084.12), and North Korea (1,927.59; 95% UI: 1,683.71–2,178.28), while the highest ASIRs were observed in Brazil (6,249.93; 95% UI: 5,569.09–6,875.72), Mexico (6,219.29; 95% UI: 5,559.67–6,822.16), and Paraguay (6,177.95; 95% UI: 5,528.53–6,806.60). National-level heterogeneity was evident in EAPC patterns: the Maldives (0.89; 95% CI: 0.56–1.22), South Korea (0.84; 95% CI: 0.55–1.13), and Iran (0.71; 95% CI: 0.48–0.93) exhibited significant increases, while the USA showed a marked decline (−0.35; 95% CI: −0.60 to −0.10) (Supplementary Table 1 and Fig. 1,2).

Trends in the prevalence of GERD in 2021

The total number of prevalent GERD cases worldwide increased from 450 million in 1990 to 830 million in 2021, while the age-standardized prevalence rate (ASPR) showed a more modest increase, from 9,516.49 per 100,000 population (95% UI: 8,427.33–10,664.72) in 1990 to 9,838.60 per 100,000 (95% UI: 8,732.46–11,056.05) in 2021. Among SDI regions, high- and high-middle-SDI regions exhibited a decline in ASPR, whereas only middle-SDI regions demonstrated an increase, with an EAPC of 0.27 (95% CI: 0.24–0.30).

In 2021, the highest ASPR was observed in Tropical Latin America (16,681.34 per 100,000 population; 95% UI: 14,832.41–18,433.23), followed by Central Latin America (16,429.42 per 100,000 population; 95% UI: 14,570.23–18,291.39), the Caribbean (16,408.86 per 100,000 population; 95% UI: 14,498.52–18,308.49), and Andean Latin America (16,405.98 per 100,000 population; 95% UI: 14,495.47–18,306.62). The highest ASPRs at the national level were observed in Paraguay (16,774.68; 95% UI: 14,835.88–18,709.59) and Brazil (16,678.10;

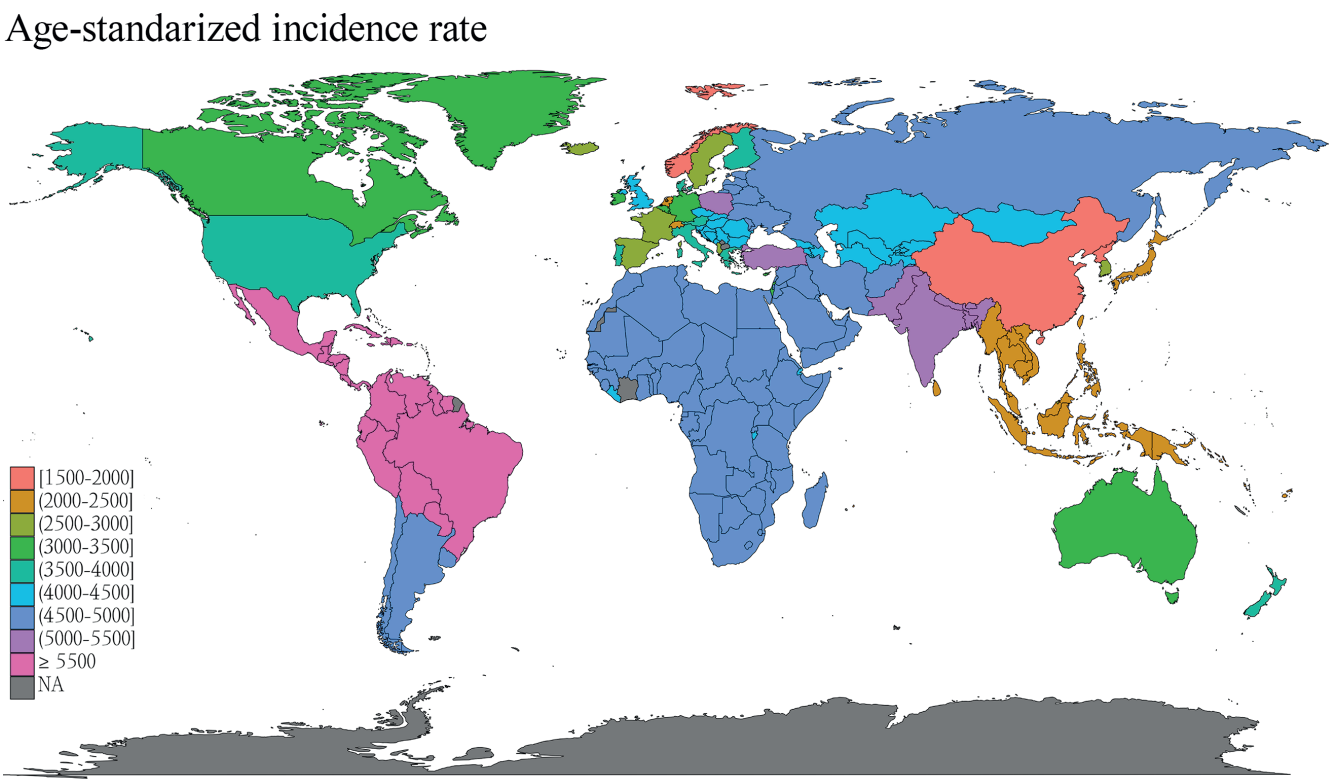


Fig. 1. Age-standardized incidence rate (ASIR) of GERD in 2021

GERD – gastroesophageal reflux disease.

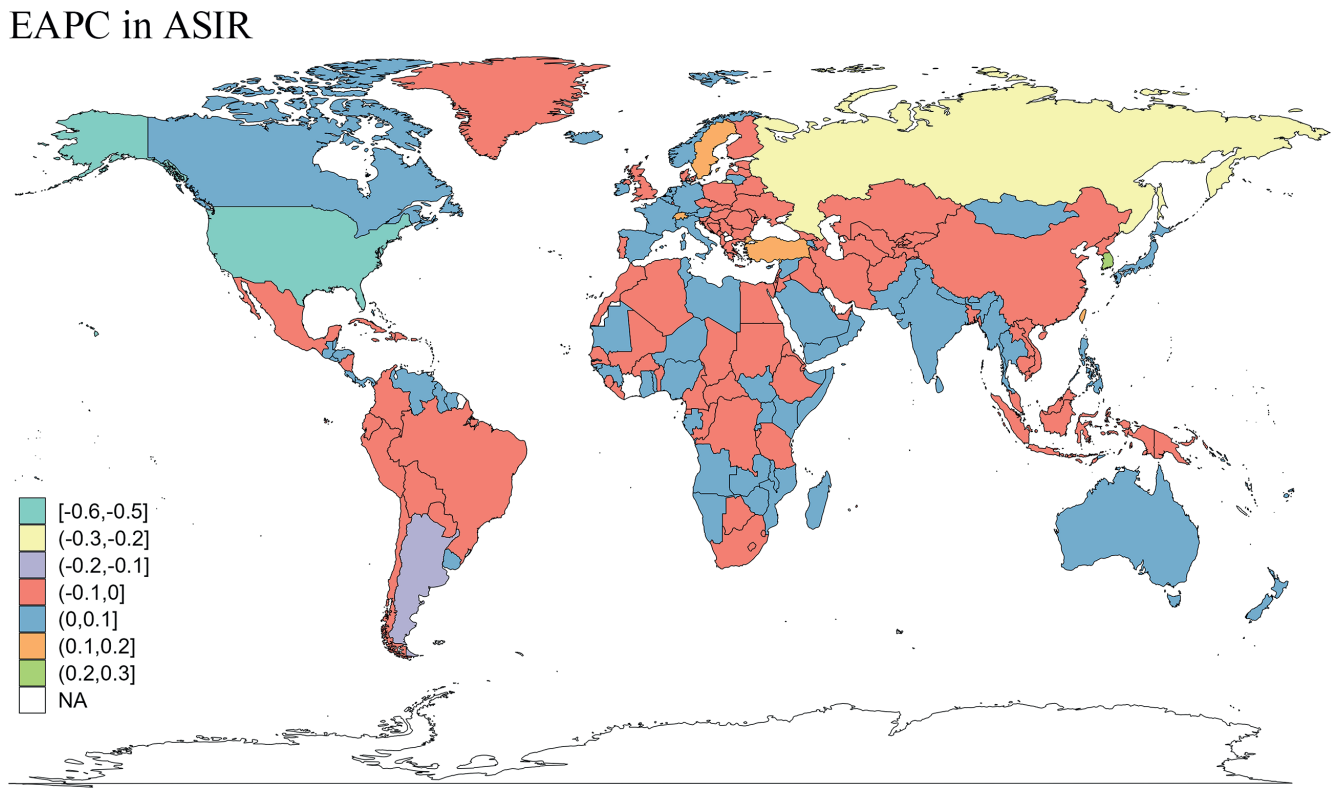


Fig. 2. Estimated annual percentage change (EAPC) in the ASIR of GERD from 1990 to 2021

ASIR – age-standardized incidence rate; GERD – gastroesophageal reflux disease.

95% UI: 14,840.12–18,432.84), whereas the lowest ASPRs were observed in China (4,540.66 per 100,000 population; 95% UI: 3,950.76–5,156.56) and Norway (4,330.94 per 100,000 population; 95% UI: 3,783.45–4,956.78). South Korea exhibited the most significant upward trend in ASPR (EAPC: 1.00; 95% CI: 0.68–1.33), while the USA showed a declining trend (EAPC: –0.49; 95% CI: –0.77 to –0.21) (Supplementary Table 2 and Fig. 3,4).

Trends in the YLDS of GERD in 2021

During the period from 1990 to 2021, the age-standardized years lived with disability (ASYLDs) rate exhibited a gradual upward trend, increasing from 73 cases per 100,000 population (95% UI: 36.75–129.66) in 1990 to 75.56 cases per 100,000 population (95% UI: 38.05–133.87) in 2021, with an EAPC of 0.37% (95% CI: 0.29–0.45). Gender-specific analysis revealed that the ASYLDs rate increased from 70.37 (95% UI: 35.3–125.34) and 75.62 (95% UI: 38.18–133.94) cases per 100,000 population in 1990 to 72.95 (95% UI: 36.6–129.73) and 78.14 (95% UI: 39.48–137.95) cases per 100,000 population in 2021 for men and women, respectively, with a more pronounced upward trend observed in men.

Geospatial analysis demonstrated heterogeneous temporal patterns across different SDI regions and geographical areas. While low-middle-SDI regions, African regions (Central, Eastern, Southern, and Western), Australia, Eastern Europe, and the Oceania region showed nonsignificant trends, high-SDI regions (–0.19; 95% CI: –0.27 to –0.10), high-middle-SDI regions (–0.23; 95% CI: –0.31 to –0.16), and North America (–0.46; 95% CI: –0.74 to –0.19) exhibited significant improvements in ASYLDs (EAPC < 0%). Conversely, Southeast Asia, South Asia, Central Latin America, High-Income Asia Pacific, and Andean Latin America showed significant increases in disease burden.

At the national level, South Korea (0.99; 95% CI: 0.67–1.31), the Maldives (0.99; 95% CI: 0.63–1.35), Iran (0.78; 95% CI: 0.53–1.03), Saint Lucia (0.74; 95% CI: 0.55–0.92), Turkey (0.74; 95% CI: 0.61–0.87), and Libya (0.73; 95% CI: 0.47–0.98) demonstrated substantial increases in GERD-related burden. In contrast, the USA showed a significant improvement trend (–0.52; 95% CI: –0.79 to –0.24) (Supplementary Table 3 and Fig. 5,6).

The trends of the incidence, prevalence, and YLDs of GERD in different gender-age groups

In the global cohort aged 0–95 years in 2021, the 35–39-year age group exhibited the highest burden of GERD. Women within this demographic reached peak values for incident cases (17,272,298 cases), prevalent cases (45,317,937 cases), and YLDs (351,061 cases). Concurrently, men in this age group reached the highest values across all age groups for incident cases (16,463,111 cases), prevalent cases (42,955,218 cases), and YLDs (337,582 cases). Notably, women consistently demonstrated higher prevalence, incidence, and YLDs than men across all age groups.

Age-specific incidence rates progressively increased, reaching an initial peak at ages 40–44 years, followed by a decline. This trend reversed after age 50, reaching a secondary peak at ages 60–64 years. After age 65, incidence rates demonstrated a sustained decline, although a resurgence was observed beyond age 89, with men exhibiting higher incidence rates than women during this geriatric phase. Age-specific prevalence peaked at ages 70–74 years, subsequently declined, and showed a secondary resurgence at ages 90–94 years. The YLD rates remained relatively stable from ages 40–84 years, peaking at ages 70–74 years. A sex-specific crossover phenomenon was observed: women maintained higher YLD rates until age 85, beyond which male rates predominated (Fig. 7).

Global prediction model for the burden of GERD from 2022 to 2050 (based on ARIMA)

The model predicts a sustained increase in incidence, prevalence, and YLDs for both sexes during this period. Specifically, the total number of incident cases is projected to exceed 200 million by 2032 among women and by 2042 among men. By 2050, the ASIR is expected to rise to 4,363 per 100,000 population for women and 3,996 per 100,000 population for men. Similarly, the number of prevalent cases among women is projected to surpass 600 million by 2043, with the ASPR reaching 11,121 per 100,000 for women and 10,135 per 100,000 for men by 2050. Furthermore, the ASYLD rate is anticipated to increase to 90 per 100,000 for women and 81 per 100,000 for men by 2050, reflecting a growing burden of disability attributable to GERD (Fig. 8).

Discussion

Our study provides the most recent estimates of the global burden of GERD across 204 countries based on the GBD 2021 study. By systematically evaluating incidence, prevalence, and YLDs at global, national, and subnational levels – stratified by SDI quintiles, sex, and age groups – we found that in 2021 there were 324 million incident cases, 826 million prevalent cases, and 6.34 million YLDs attributable to GERD worldwide. The burden exhibited significant geographical heterogeneity: ASIRs and ASPRs were highest in Tropical Latin America and lowest in East Asia, a disparity potentially driven by regional variations in dietary patterns, obesity prevalence, alcohol and tobacco consumption, and population aging.[3,14] Meanwhile, socioeconomic disparities markedly influenced GERD burden, with low- to middle-SDI regions exhibiting the highest ASIRs, ASPRs, and ASYLDs compared with other SDI

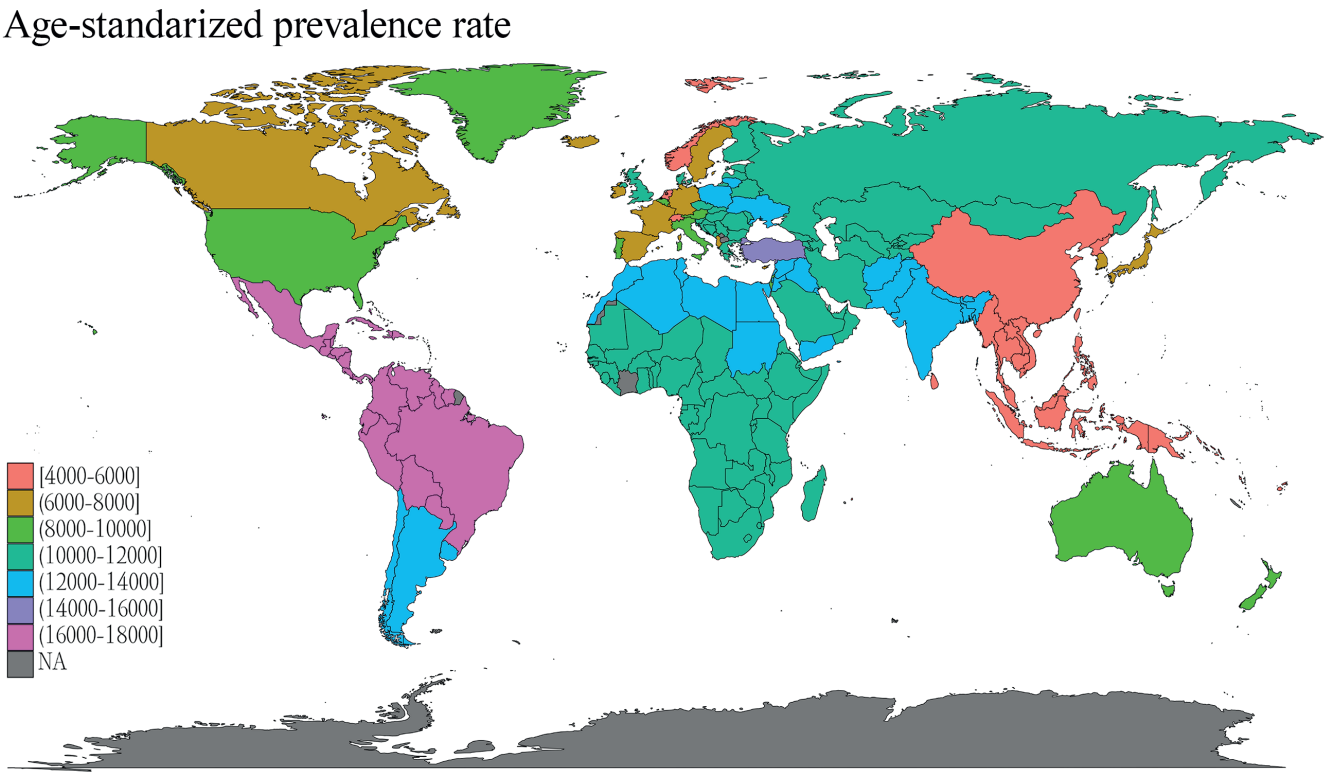


Fig. 3. Age-standardized prevalence rate (ASPR) of GERD in 2021

GERD – gastroesophageal reflux disease.

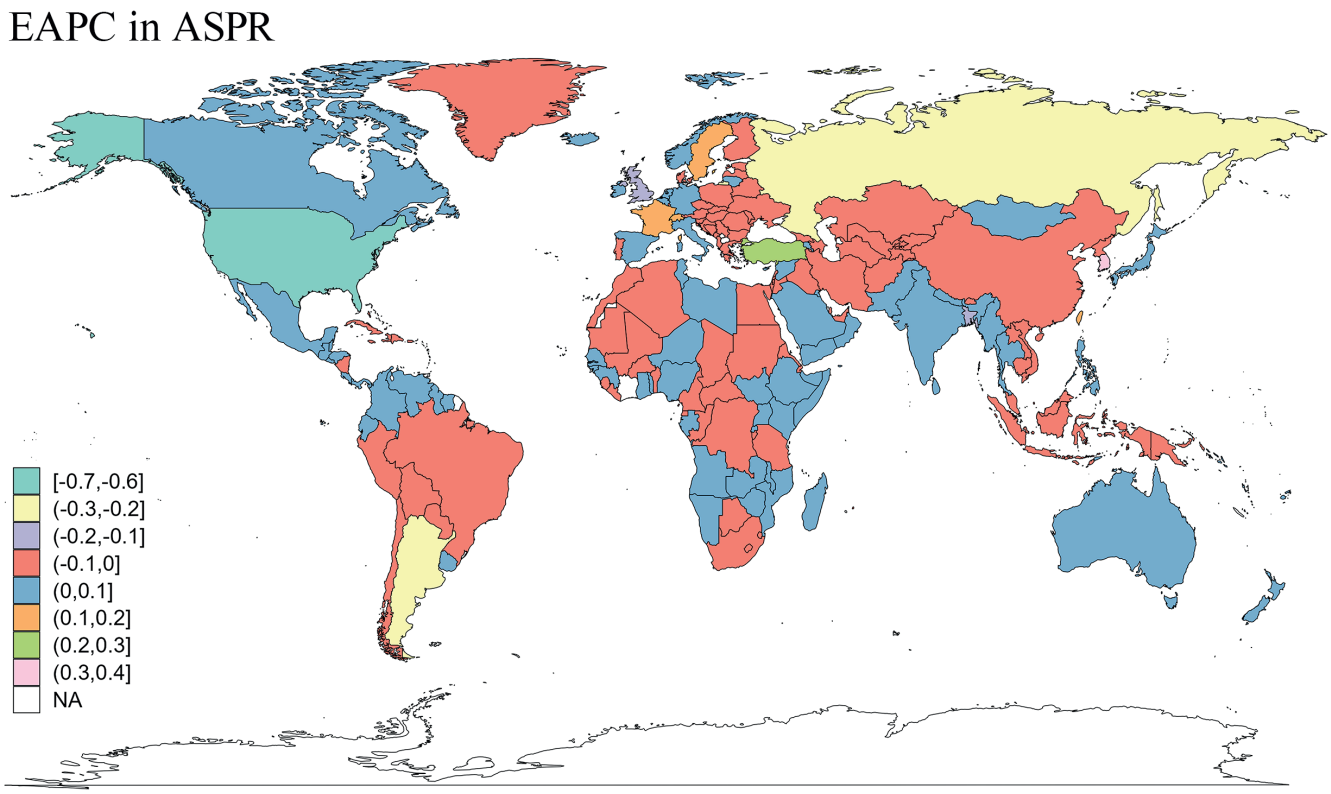


Fig. 4. Estimated annual percentage change (EAPC) in the ASPR of GERD from 1990 to 2021

ASPR – age-standardized prevalence rate; GERD – gastroesophageal reflux disease.

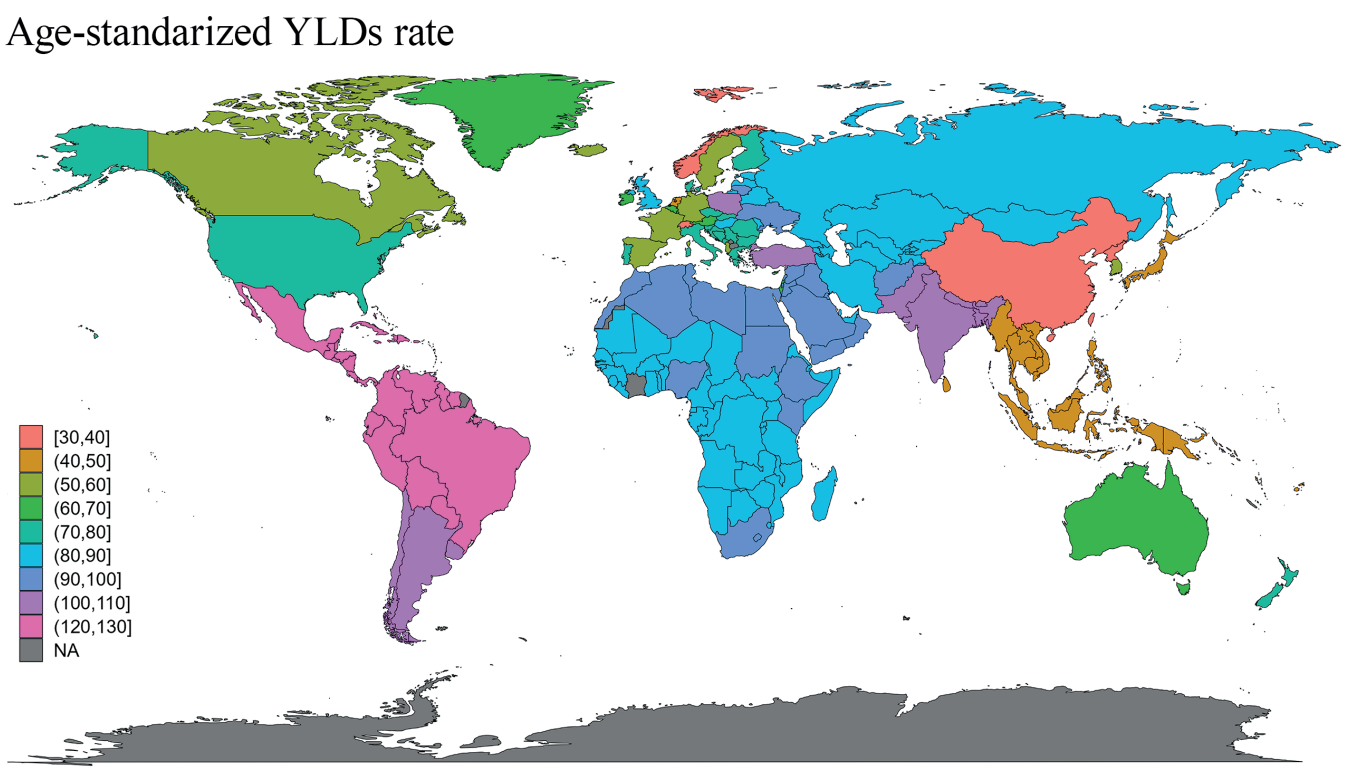


Fig. 5. Age-standardized years lived with disability (ASYLDs) due to GERD in 2021

GERD – gastroesophageal reflux disease.

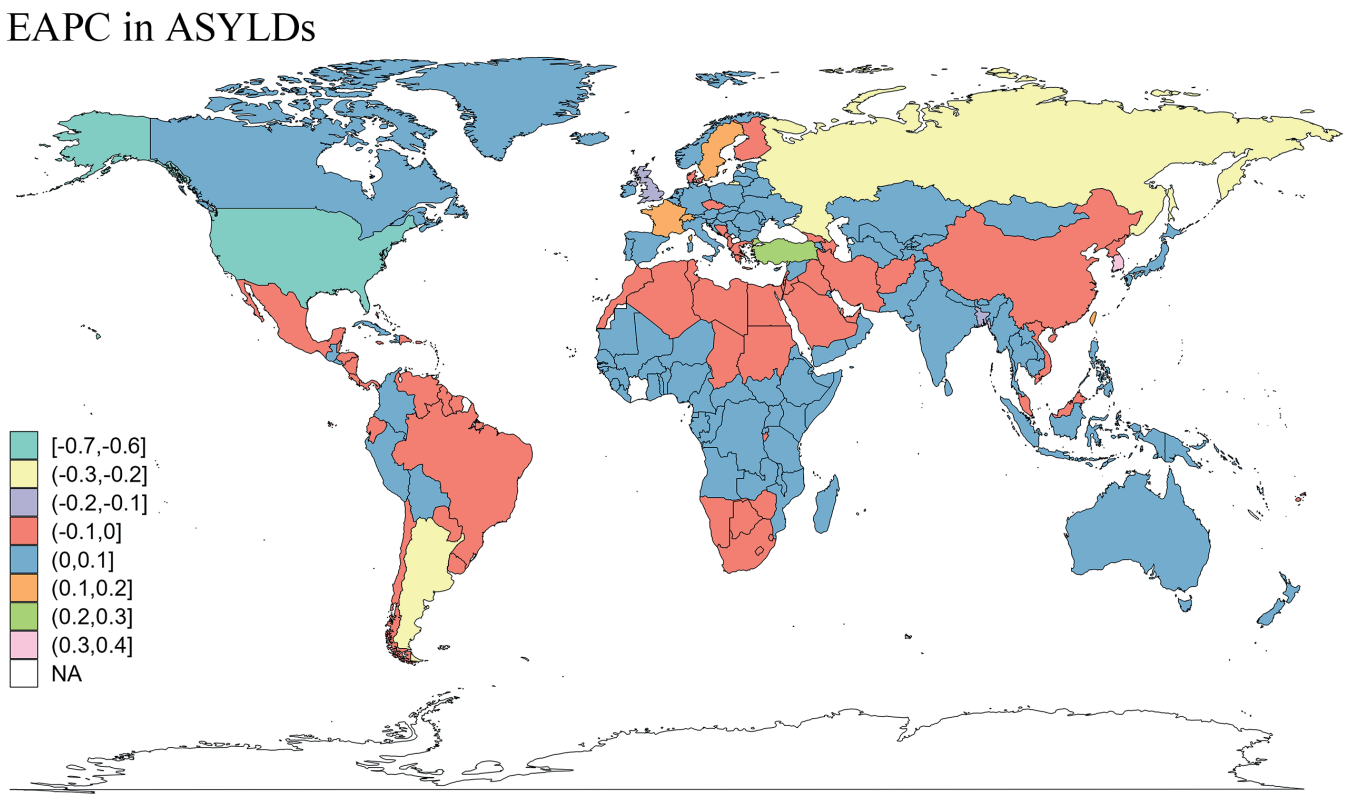


Fig. 6. Estimated annual percentage change (EAPC) in ASYLDs due to GERD from 1990 to 2021

ASYLDs – age-standardized years lived with disability; GERD – gastroesophageal reflux disease.

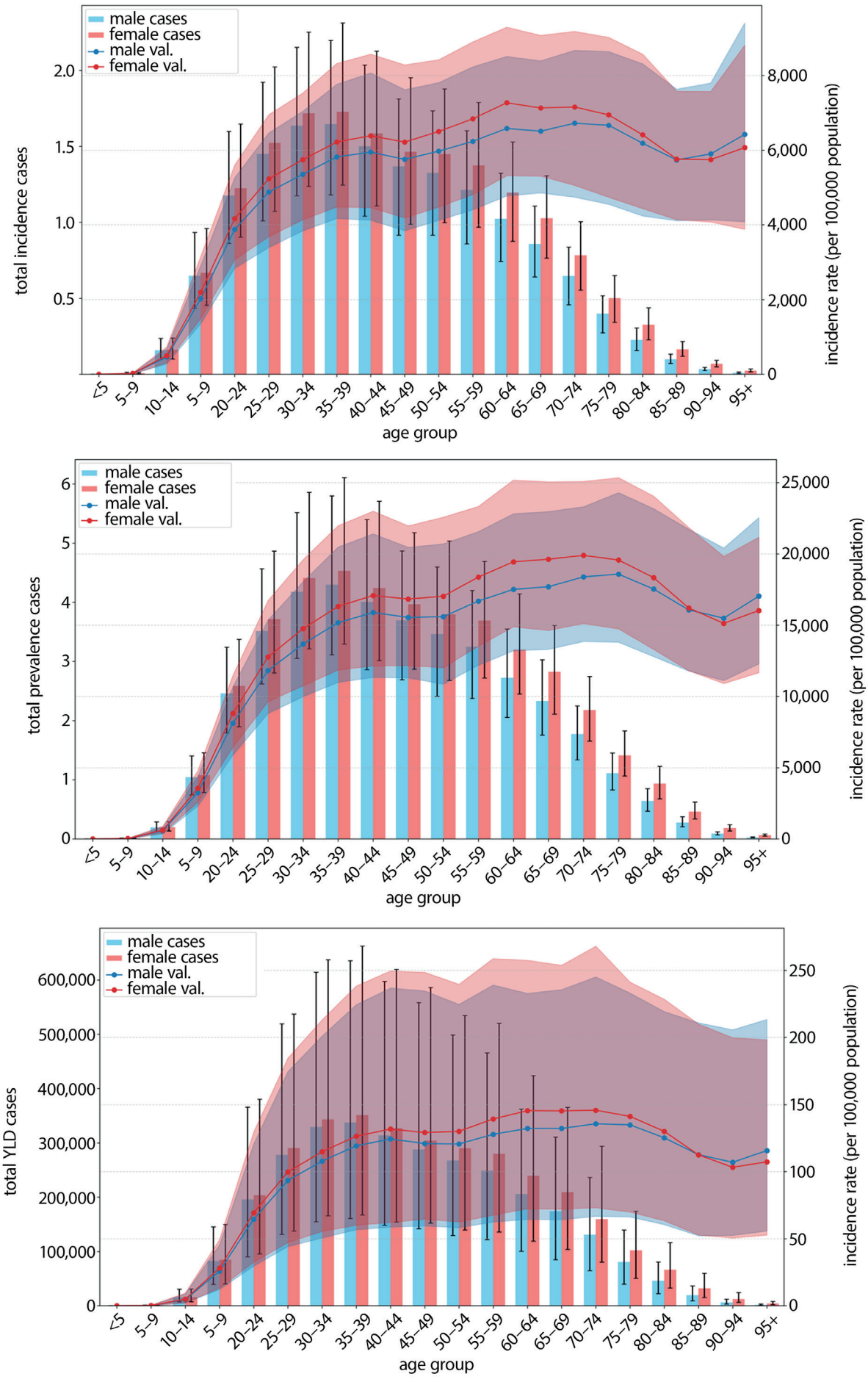


Fig. 7. Age patterns by sex for the total number and age-specific prevalence rate (ASPR), age-specific incidence rate (ASID), and age-specific YLD rate due to GERD at the global level in 2021. Error bars indicate the 95% UI for the number of cases. Shading indicates the 95% UI for the rates

YLD – year lived with disability; GERD – gastroesophageal reflux disease; UI – uncertainty intervals.

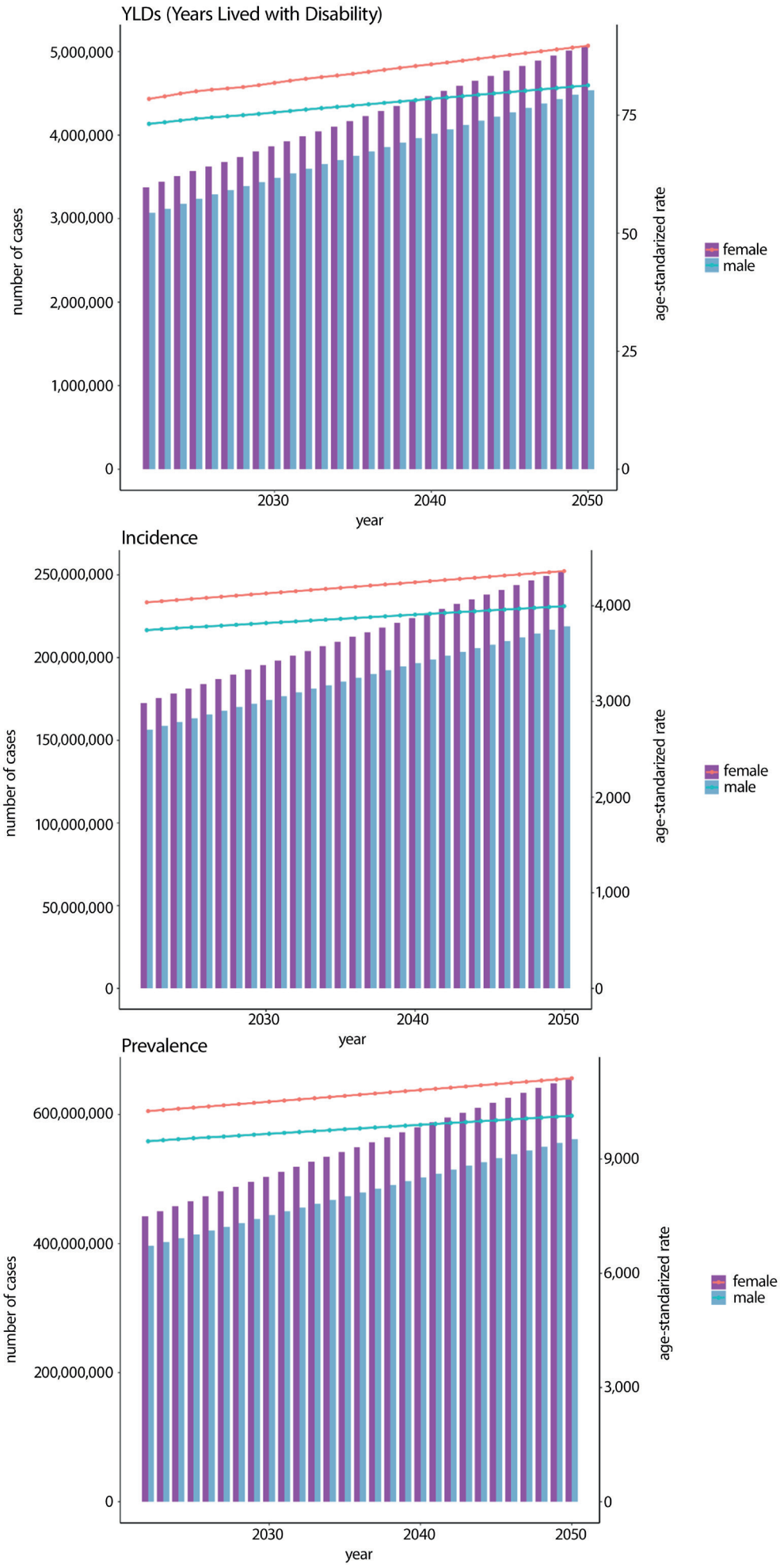


Fig. 8. Predictions of the global incidence, prevalence, and YLDs of GERD from 2022 to 2050

YLDs – years lived with disability; GERD – gastroesophageal reflux disease.

quintiles, likely reflecting inequities in healthcare access, diagnostic delays, and suboptimal management of chronic conditions. Furthermore, common digestive disorders such as celiac disease may present with symptoms overlapping those of GERD, including abdominal discomfort and reflux-like sensations, which can complicate differential diagnosis and potentially lead to misdiagnosis of GERD or overestimation of its incidence in certain populations.[15] Additionally, NSAIDs and corticosteroids have been shown to exacerbate GERD by impairing esophageal sphincter function, reducing mucosal protection, and increasing gastric acid secretion. The risk of reflux symptoms increases one-to-twofold among long-term users. These underrecognized factors may contribute to the continuous rise in GERD incidence, and the projected burden is expected to increase further by 2050. This highlights the need for targeted preventive measures. Future research should integrate detailed data on comorbidities and pharmacological treatments to improve prediction accuracy.

Our analysis yielded the following principal findings. First, the ASIR, ASPR, and ASYLDs associated with GERD exhibited a consistent upward trajectory globally from 1990 to 2021, underscoring the escalating disease burden. These epidemiological patterns align with findings from previous studies published in 2017 and 2019, confirming the persistent temporal progression of GERD-related health impacts. Geospatial analysis revealed distinct regional variations, with the highest ASIR estimates observed in Brazil, Mexico, and Paraguay, while ASPR peaked in Paraguay, Brazil, and El Salvador.

Notably, divergent temporal patterns emerged between 2019 and 2021, with Asia and High-Income Asia Pacific regions demonstrating accelerated growth rates in disease metrics, in contrast to the declining trend observed in High-Income North America. This epidemiological divergence suggests that, beyond universal drivers such as population growth and aging, region-specific determinants – including dietary patterns, obesity prevalence, and healthcare accessibility – may modulate GERD burden dynamics. Numerous studies have also shown that obesity and overweight increase intra-abdominal pressure, weaken the lower esophageal sphincter, and exacerbate reflux, thereby directly contributing to increased GERD incidence.[16,17]

Furthermore, we identified a significant disparity in disease burden, with disproportionately higher impacts observed in socioeconomically disadvantaged regions characterized by limited healthcare resource allocation and constrained diagnostic capacity, which is consistent with previous findings.[9,18] Finally, projection modeling using ARIMA algorithms predicts a sustained global increase in both prevalence rates and incidence counts from 2022 to 2050, signaling an urgent need for targeted public health interventions.

First, greater emphasis should be placed on promoting healthy dietary patterns, smoking and alcohol control policies, and obesity management strategies in high-burden regions such as Tropical Latin America. Second, healthcare capacity at the primary care level in low- and middle-SDI regions should be strengthened through the implementation of low-cost diagnostic tools and improved access to treatment. Optimization of healthcare resource allocation and expansion of telemedicine and digital health tools may improve diagnostic accessibility in remote areas, shorten diagnostic delays, and improve chronic reflux management, thereby reducing progression to severe complications. Third, regional epidemiological surveillance and dynamic monitoring systems should be strengthened to track changes in disease burden. In regions with rapidly increasing rates, such as Asia and High-Income Asia Pacific, GERD prevention strategies should be integrated into primary healthcare systems, with emphasis on early symptom recognition and lifestyle counseling. In regions with declining trends, such as High-Income North America, best practices in healthcare accessibility and standardized management could be shared to guide interventions in other regions. Finally, these measures should align with international GERD guidelines while being adapted to local socioeconomic conditions to ensure feasibility and equity across populations with different levels of socioeconomic development, thereby reducing the burden of GERD and narrowing regional disparities.

When assessing the severity of GERD, several high-risk factors related to the occurrence and progression of the disease, such as age, race, and sex, should be taken into account. Our research found that the 35–39-year age group exhibited the highest burden of GERD. Previous studies have shown that psychological factors play an important role in the pathogenesis of GERD.[19] Psychological stress can increase both the perception of heartburn and the severity of GERD symptoms.[20] Fass et al. indicated that acute stress can enhance esophageal sensitivity to acid exposure in patients with reflux esophagitis or GERD.[21] Individuals aged 35–39 often face multiple stressors, including career advancement, mortgage obligations, and raising children. Emotional conditions such as anxiety and depression may lead to irregular eating habits, overeating, and increased smoking and alcohol consumption, thereby increasing the incidence of GERD.

This study found that the prevalence of GERD peaked at ages 70–74 years and subsequently declined. The decline in prevalence after age 75 may reflect the combined effects of "selective survival bias" and "physiological or diagnostic changes."[22] First, GERD is closely associated with several risk factors, such as severe obesity, smoking, and poorly controlled comorbidities, which also increase mortality related to cardiovascular disease (CVD), cancer, and other age-related conditions. Patients with these high-risk characteristics may die before the age of 75, resulting in a "pseudo-decline" in GERD prevalence among individuals surviving to older ages. Second, aging-related

reductions in esophageal sensory perception and sensitivity may decrease symptom severity or awareness in elderly individuals, leading to underreporting of GERD diagnoses. Third, individuals older than 75 years often experience multiple comorbidities, such as CVD and dementia, which may divert healthcare attention and contribute to underdiagnosis of GERD symptoms in epidemiological data.

To date, research on the incidence and prevalence of GERD remains limited. For instance, a systematic review by Singendonk et al. estimated that GERD affects approx. 23–40% of pediatric populations.[23] In contrast to earlier findings, the GBD 2019 database and a study by Nirwan et al. identified China as having the lowest ASPR for GERD, whereas the latest GBD 2021 data indicate that Norway now has the lowest ASPR.[24,25] Notably, our analysis revealed significant improvements in ASIRs, ASPRs, and ASYLDs in North America. Consistent with previous research, the burden of GERD is closely associated with the SDI. Although high-fat and high-sugar diets in regions such as North America may elevate obesity rates and consequently increase GERD risk, greater health awareness, adequate healthcare resources, and standardized diagnostic and monitoring protocols appear to effectively mitigate GERD incidence.[26,27]

Limitations of the study

Our study has several limitations, the most significant of which is the scarcity of data in many regions. Although the analysis of GERD incorporated prevalence and incidence data from a total of 204 location-years, this is substantially fewer than the data available for other chronic conditions such as diabetes and coronary heart disease. The limited availability of data constrains the precision of regional burden estimates, particularly in understudied or data-deficient areas. To address this gap, future iterations of the GBD study should prioritize inclusion of additional high-quality data sources, with a specific focus on underrepresented regions.

Meanwhile, GERD is a heterogeneous disease with a wide range of symptoms and complications. The incidence estimates in this study reflect "clinically diagnosed GERD"; however, GERD may be underreported because of the presence of atypical manifestations. Future research should therefore consider including atypical symptoms and complications of GERD, such as erosive esophagitis, BE, and stenosis, as well as hospitalization rates related to complications and proton pump inhibitor (PPI) use or prescription as supplementary indicators. These revisions may enhance the transparency, scope, and clinical relevance of the analyzed data. Furthermore, there is a critical need for more granular data on GERD incidence, prevalence, severity, and frequency across different age groups, particularly in regions where such data are currently sparse or absent.

Conclusions

Our study demonstrates a substantial increase in the global burden of GERD from 1990 to 2021. Projections indicate that this burden will continue to increase from 2022 to 2050, imposing substantial strain on healthcare systems worldwide, especially among female populations. Middle- and lower-middle-SDI regions consistently exhibit disproportionately high GERD burdens compared with high-income regions. To address these challenges, future efforts must prioritize implementation of precision diagnostic protocols and evidence-based interventions tailored to regional epidemiological profiles. Multifaceted strategies, including enhanced surveillance systems, equitable access to PPIs, and community-based education regarding lifestyle modifications, are critical to mitigating GERD-related complications and reducing the growing impact of GERD on global health equity.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17132213. The package contains the following files:

Supplementary Table 1. Incidence of GERD in 2021 for both sexes and all locations, with EAPC from 1990 and 2021.

Supplementary Table 2. Prevalence of GERD in 2021 for both sexes and all locations, with EAPC from 1990 and 2021.

Supplementary Table 3. YLDs of GERD in 2021 for both sexes and all locations, with EAPC from 1990 and 2021.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Kejin Li https://orcid.org/0009-0009-2460-4363
Hui Li https://orcid.org/0000-0003-4742-8912
Zhentao An https://orcid.org/0000-0002-4977-3093
Xiangxiang Xu https://orcid.org/0009-0007-3420-6655
Jing Zuo https://orcid.org/0009-0004-0794-2823
Jiaxin Li https://orcid.org/0009-0003-9997-228X
Xueli Qian https://orcid.org/0009-0009-7177-2941
Liu Liu https://orcid.org/0009-0003-5541-9503
Jingjing Cui https://orcid.org/0009-0005-3690-2289

References

1. Rosen R, Vandenplas Y, Singendonk M, et al. Pediatric Gastroesophageal Reflux Clinical Practice Guidelines: Joint Recommendations of the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition. *J Pediatr Gastroenterol Nutr*. 2018;66(3):516–554. doi:10.1097/MPG.0000000000001889
2. Mousa H, Hassan M. Gastroesophageal reflux disease. *Pediatr Clin North Am*. 2017;64(3):487–505. doi:10.1016/j.pcl.2017.01.003
3. Fass R, Boeckxstaens GE, El-Serag H, Rosen R, Sifrim D, Vaezi MF. Gastro-oesophageal reflux disease. *Nat Rev Dis Primers*. 2021;7(1):55. doi:10.1038/s41572-021-00287-w

4. Mohy-ud-din N, Krill TS, Shah AR, et al. Barrett's esophagus: What do we need to know? *Disease-a-Month*. 2020;66(1):100850. doi:10.1016/j.disamonth.2019.02.003
5. Weusten BLAM, Bisschops R, Dinis-Ribeiro M, et al. Diagnosis and management of Barrett esophagus: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy*. 2023;55(12):1124–1146. doi:10.1055/a-2176-2440
6. Eusebi LH, Ratnakumaran R, Yuan Y, Solaymani-Dodaran M, Bazzoli F, Ford AC. Global prevalence of, and risk factors for, gastro-oesophageal reflux symptoms: A meta-analysis. *Gut*. 2018;67(3):430–440. doi:10.1136/gutjnl-2016-313589
7. Hallan A, Bomme M, Hveem K, Møller-Hansen J, Ness-Jensen E. Risk factors on the development of new-onset gastroesophageal reflux symptoms. A population-based prospective cohort study: The HUNT Study. *Am J Gastroenterol*. 2015;110(3):393–400. doi:10.1038/ajg.2015.18
8. Dirac MA, Safiri S, Tsoi D, et al. The global, regional, and national burden of gastro-oesophageal reflux disease in 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol*. 2020;5(6):561–581. doi:10.1016/S2468-1253(19)30408-X
9. Li N, Yang WL, Cai MH, et al. Burden of gastroesophageal reflux disease in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *BMC Public Health*. 2023;23(1):582. doi:10.1186/s12889-023-15272-z
10. Wei Y, Liu E, Peng J, Liu Y, Sun X, Yao X. Global burden of esophageal diseases: A comprehensive analysis of disease trends and risk factors from 1990 to 2021. *BMC Gastroenterol*. 2025;25(1):528. doi:10.1186/s12876-025-03988-8
11. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204–1222. doi:10.1016/S0140-6736(20)30925-9
12. Shen Z, Luo H. The impact of schistosomiasis on the Global Disease Burden: A systematic analysis based on the 2021 Global Burden of Disease Study. *Parasite*. 2025;32:12. doi:10.1051/parasite/2025005
13. Shumway RH, Stoffer DS. *Time Series Analysis and Its Applications: With R Examples*. 4th ed. Cham, Switzerland: Springer; 2017. ISBN:978-3-319-52452-8.
14. Ness-Jensen E, Lagergren J. Tobacco smoking, alcohol consumption and gastro-oesophageal reflux disease. *Best Pract Res Clin Gastroenterol*. 2017;31(5):501–508. doi:10.1016/j.bpg.2017.09.004
15. Raiteri A, Granito A, Giamperoli A, Catenaro T, Negrini G, Tovoli F. Current guidelines for the management of celiac disease: A systematic review with comparative analysis. *World J Gastroenterol*. 2022;28(1):154–176. doi:10.3748/wjg.v28.i1.154
16. Barba Orozco E, Ezquerra Duran A. Study of gastroesophageal reflux: Interpretation of functional tests in the obese patient. *Cir Esp (Engl Ed)*. 2023;101(Suppl 4):S8–S18. doi:10.1016/j.cireng.2023.01.015
17. Alimi Y, Azagury DE. Gastroesophageal reflux disease and the patient with obesity. *Gastroenterol Clin North Am*. 2021;50(4):859–870. doi:10.1016/j.gtc.2021.08.010
18. Bai Z, Wang H, Shen C, An J, Yang Z, Mo X. The global, regional, and national patterns of change in the burden of nonmalignant upper gastrointestinal diseases from 1990 to 2019 and the forecast for the next decade. *Int J Surg*. 2025;111(1):80–92. doi:10.1097/JS9.0000000000001902
19. Hartono JL, Mahadeva S, Goh K. Anxiety and depression in various functional gastrointestinal disorders: Do differences exist? *J Digest Dis*. 2012;13(5):252–257. doi:10.1111/j.1751-2980.2012.00581.x
20. Lee SP, Sung IK, Kim JH, Lee SY, Park HS, Shim CS. The effect of emotional stress and depression on the prevalence of digestive diseases. *J Neurogastroenterol Motil*. 2015;21(2):273–282. doi:10.5056/jnm14116
21. Fass R, Naliboff BD, Fass SS, et al. The effect of auditory stress on perception of intraesophageal acid in patients with gastroesophageal reflux disease. *Gastroenterology*. 2008;134(3):696–705. doi:10.1053/j.gastro.2007.12.010
22. Becher A, Dent J. Systematic review: Ageing and gastro-oesophageal reflux disease symptoms, oesophageal function and reflux oesophagitis: Systematic review: ageing and GERD. *Aliment Pharmacol Ther*. 2011;33(4):442–454. doi:10.1111/j.1365-2036.2010.04542.x
23. Singendonk M, Goudswaard E, Langendam M, et al. Prevalence of gastroesophageal reflux disease symptoms in infants and children: A systematic review. *J Pediatr Gastroenterol Nutr*. 2019;68(6):811–817. doi:10.1097/MPG.0000000000002280
24. Zhang D, Liu S, Li Z, Wang R. Global, regional and national burden of gastroesophageal reflux disease, 1990–2019: Update from the GBD 2019 study. *Ann Med*. 2022;54(1):1372–1384. doi:10.1080/07853890.2022.2074535
25. Nirwan JS, Hasan SS, Babar ZUD, Conway BR, Ghori MU. Global prevalence and risk factors of gastro-oesophageal reflux disease (GORD): Systematic review with neta-analysis. *Sci Rep*. 2020;10(1):5814. doi:10.1038/s41598-020-62795-1
26. Katzka DA, Kahrilas PJ. Advances in the diagnosis and management of gastroesophageal reflux disease. *BMJ*. 2020;371:m3786. doi:10.1136/bmj.m3786
27. Su B, Dunst C, Gould J, et al. Experience-based expert consensus on the intra-operative usage of the Endoflip impedance planimetry system. *Surg Endosc*. 2021;35(6):2731–2742. doi:10.1007/s00464-020-07704-3

SGLT2 inhibitors in metabolic dysfunction-associated steatotic liver disease (MASLD/NAFLD): Mechanisms, clinical benefits, and outcomes

Jan Paleček[1,2,D], Petr Hříbek[1,2,E,F], Denisa Janíčková Ždarská[1,A], Marie Nováková[3,B], Pavel Skořepa[4,2,E], Petr Urbánek[1,2,E]

[1] Department of Medicine, 1st Faculty of Medicine Charles University, Military University Hospital Prague, Czech Republic
[2] Department of Military Internal Medicine and Military Hygiene, Military Faculty of Medicine, University of Defense, Hradec Kralove, Czech Republic
[3] Department of Pathology, Military University Hospital Prague, Czech Republic
[4] 3rd Department of Internal Medicine-Metabolism and Gerontology, University Hospital Hradec Kralove, Czech Republic

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1085–1092

Address for correspondence
Petr Hříbek
E-mail: Petr.Hribek@uvn.cz

Funding sources
None declared

Conflict of interest
None declared

Received on June 13, 2025
Reviewed on July 18, 2025
Accepted on August 22, 2025

Published online on June 29, 2026

Abstract

Metabolic dysfunction-associated steatohepatitis (MASH) represents a progressive form of metabolic dysfunction-associated steatotic liver disease (MASLD) and contributes significantly to morbidity and mortality in developed countries. Current clinical practice primarily relies on dietary and lifestyle modifications, while specialized obesity management is used less frequently. However, the long-term efficacy of these approaches remains unsatisfactory. Among pharmacological options, gliflozins have emerged as a promising class of agents with potential benefits in reducing histological activity and improving the prognosis of MASLD/MASH. The beneficial effects of gliflozins in diabetes, cardiology, and nephrology have already been demonstrated, and their use is becoming increasingly common across various clinical indications. However, their role in the treatment of MASLD/MASH has not yet been fully clarified. Preliminary studies suggest that gliflozins may positively influence disease progression in these conditions. This article provides a comprehensive overview of MASLD/MASH and reviews current evidence regarding the therapeutic effects of gliflozins on these liver diseases.

Key words: non-alcoholic fatty liver disease (NAFLD), metabolic dysfunction-associated steatohepatitis (MASH), metabolic dysfunction-associated steatotic liver disease (MASLD), sodium-glucose transporter 2 inhibitors (SGLT2 inhibitors), type 2 diabetes mellitus (T2DM)

Cite as
Paleček J, Hříbek P, Janíčková Ždarská D, Nováková M, Skořepa P, Urbánek P. SGLT2 inhibitors in metabolic dysfunction-associated steatotic liver disease (MASLD/NAFLD): Mechanisms, clinical benefits, and outcomes. *Adv Clin Exp Med*. 2026;35(6):1085–1092. doi:10.17219/acem/209813

DOI
10.17219/acem/209813

Highlights

- With the rising use of sodium-glucose cotransporter-2 inhibitors (SGLT2i), their effects on metabolic dysfunction-associated steatotic liver disease (MASLD) and metabolic dysfunction-associated steatohepatitis (MASH) should be established.
- Available data suggest a potential beneficial effect of SGLT2i in improving the course of MASLD/MASH.
- Indirect assessment methods suggest a possible effect of SGLT2i on fibrosis, steatosis, and oxidative stress; however, histological verification is still required.

Introduction

Metabolic dysfunction-associated steatotic liver disease (MASLD) is currently the most common liver disease worldwide. Available estimates suggest that up to 1/3 of the global population may be affected by this condition. Metabolic dysfunction-associated steatotic liver disease is considered the hepatic component of metabolic syndrome and contributes significantly to morbidity and mortality in developed countries.[1] The diagnostic criteria include hepatic steatosis detected through imaging and/or histology, together with the presence of at least 1 of 5 cardiometabolic risk factors (Table 1).[2] The change in nomenclature and introduction of the term MASLD are intended to place greater emphasis on metabolic dysregulation, such as type 2 diabetes mellitus (T2DM) or obesity, compared with the former term non-alcoholic fatty liver disease (NAFLD). At the same time, the new nomenclature acknowledges the synergistic effect of alcohol consumption. This change in terminology reflects key findings regarding hepatic and extrahepatic mortality risk, disease associations, and the identification of high-risk individuals.[3]

Recent studies suggest that excess body weight and metabolic dysfunction independently affect the development of MASLD and cardiometabolic outcomes. In clinical practice, MASLD is commonly associated with overweight/obesity and T2DM, with more than 70% of patients with T2DM having some form of MASLD. However, studies suggest that even patients with metabolically healthy obesity, defined as the absence of metabolic disorders and/or cardiovascular disease (CVD), remain at high risk of developing significant hepatic fibrosis. Lean individuals are also not protected against MASLD development, and no difference in histological severity has been observed between patients with a body mass index (BMI) <23 kg/m^2 and those with a BMI > 25 kg/m^2. These findings indicate that both metabolic health and metabolic dysfunction, which may vary throughout a patient's life, are important factors to consider.[3,4] Metabolic dysfunction-associated steatotic liver disease is further subdivided into metabolic dysfunction-associated steatosis (MASL) and metabolic dysfunction-associated steatohepatitis (MASH). These subgroups differ substantially in clinical course and prognosis. While MASL is more commonly observed as a concomitant condition in patients with metabolic syndrome, MASH represents a progressive form that may eventually lead to hepatocyte apoptosis, fibrosis, loss of functional liver tissue, and ultimately liver cirrhosis with its associated complications.[1]

In current clinical practice, the Fibrosis-4 (FIB-4) score (age, aspartate transaminase (AST), alanine transaminase / lanine aminotransferase (ALT), and platelet count) is used to screen at-risk patients for liver fibrosis.[5]

$$FIB\text{-}4 = \frac{Age(years) \times AST\ level(U/L)}{Platelet\ count\left(10^9/L\right) \times \sqrt{ALT\ level(U/L)}}$$

If the FIB-4 value is between 1.3 and 2.6, noninvasive liver stiffness measurement (LSM) is recommended.[5,6] If the score exceeds 2.6, direct referral to a hepatogastroenterology specialist is indicated. The diagnosis of MASH is based on histological examination of liver tissue.[6,7]

Objectives

With the rising use of SGLT2i, their effects on MASLD/MASH should be established. This short review provides

Table 1. Cardiometabolic risk factors[2]

Cardiometabolic risk factors
1. BMI ≥ 25 kg/m^2 or waist circumference >93 cm in men and >80 cm in women
2. Fasting serum glucose ≥5.6 mmol/L or post-load serum glucose ≥7.8 mmol/L or glycated hemoglobin ≥39 mmol/mol or diagnosis of T2DM or active antidiabetic treatment
3. Blood pressure ≥130/85 mm Hg or targeted treatment for hypertension
4. Plasma triacylglycerol level ≥1.7 mmol/L or targeted treatment for dyslipidemia
5. Plasma HDL cholesterol ≤1 mmol/L in men, 1.3 mmol/L in women or treatment of dyslipidemia

BMI – body mass index; T2DM – type 2 diabetes mellitus; HDL – high-density lipoprotein.

an overview of the topic and attempts to link original studies and recommendations from professional societies with recent studies and meta-analyses, without using a specific sampling protocol. The aim of this article is to provide a brief introduction to the issue, promote further research in this area, and indicate where it is most needed.

Pathogenesis of MASH

The exact pathogenesis of MASH and the mechanisms underlying its progression have not yet been fully elucidated, and a multifactorial etiology is assumed. Increased concentrations of fatty acids accumulated in hepatic adipose tissue lead to oxidative stress, mitochondrial damage, hepatocyte injury, and, consequently, the development of insulin resistance. Furthermore, insulin resistance contributes to macrophage activation and the development of inflammation. The roles of genetics, epigenetics, and the gut microbiome remain incompletely understood. Diagnosis is additionally complicated by the difficulty in accurately assessing the quantity and frequency of alcohol consumption.

There is growing evidence that a high-fat diet, insulin resistance, obesity, dysregulation of peripheral lipolysis, and other metabolic risk factors contribute to increased hepatic deposition of free fatty acids. Progressive accumulation of triacylglycerols within hepatocyte cytoplasm leads to steatosis that is not attributable to excessive alcohol consumption or drug-induced liver injury. Repeated liver injury, endoplasmic reticulum oxidative stress, reactive oxygen species (ROS), and mitochondrial dysfunction subsequently promote inflammation. Insulin resistance further contributes to de novo lipogenesis and adipose tissue dysfunction, resulting in the release of inflammatory adipokines and cytokines, including tumor necrosis factor alpha (TNF-α).[8]

The development of hepatic inflammation distinguishes MASLD from MASH, as illustrated in Fig. 1, which demonstrates simple steatosis on hematoxylin and eosin (H&E) staining, in contrast to MASH shown in Fig. 2. These and other factors disrupt hepatic homeostasis and activate immune responses, leading to hepatocyte apoptosis and liver fibrosis.[9] Liver cirrhosis is frequently complicated by the development of hepatocellular carcinoma (HCC; Fig. 3).[10] Hepatocellular carcinoma is the 6th most commonly diagnosed cancer and the 3rd leading cause of cancer-related death worldwide, with a higher incidence in men. It accounts for 75–85% of primary liver cancers. MASH is considered one of the major risk factors for HCC, alongside hepatitis B virus infection, hepatitis C virus infection, and excessive alcohol consumption (>30 g/day in men and >20 g/day in women).[11]

The progression of MASLD was evaluated in a study including 718 patients who underwent more than 2 non-targeted liver biopsies performed at least 6 months apart. At the initial biopsy, 497 patients (69.2%) had MASLD, 90 (12.5%) had non-fibrotic MASH, and 131 (18.2%) had non-cirrhotic fibrosis. Over a median follow-up period of 3.4 years between biopsies, 30.4% (218/718) of patients experienced disease progression, including 12.5% (62/497) with incident non-fibrotic MASH, 24.0% (141/587) with incident fibrosis, and 5.6% (40/718) with cirrhosis.[12]

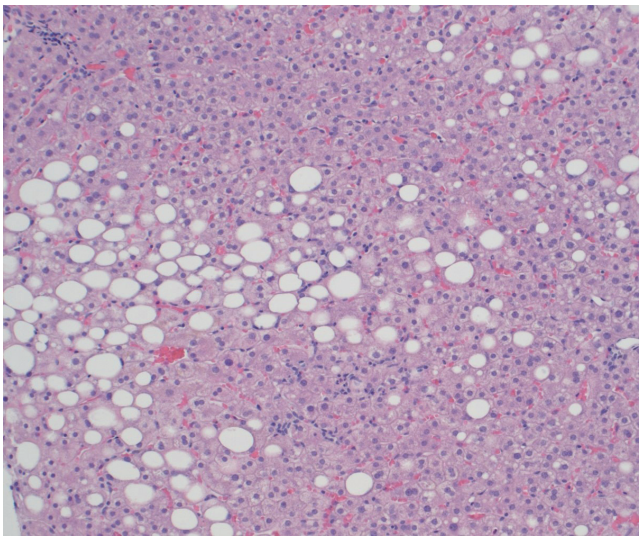


Fig. 1. Simple steatosis. Hematoxylin and eosin (H&E) staining – predominantly macrovesicular steatosis is seen without lobular necroinflammatory or ballooning changes, with no evidence of Mallory–Denk body formation

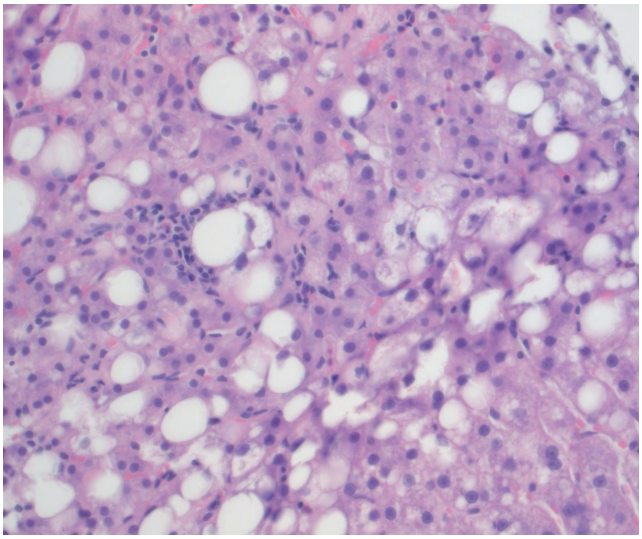


Fig. 2. Metabolic dysfunction-associated steatohepatitis (MASH). Hematoxylin and eosin (H&E) stain – predominantly macrovesicular steatosis, few hepatocytes show ballooning change and there is patchy lobular lymphocytic inflammation in association with focal hepatocytic injury

Epidemiology and current therapy of MASLD

The worldwide prevalence of MASLD is reported to be approx. 30%, while the global prevalence of MASH is estimated at 5.27%.[1] According to the Global Burden of Disease study, MASLD is the disease with the fastest-growing proportion among chronic liver diseases, including liver cirrhosis

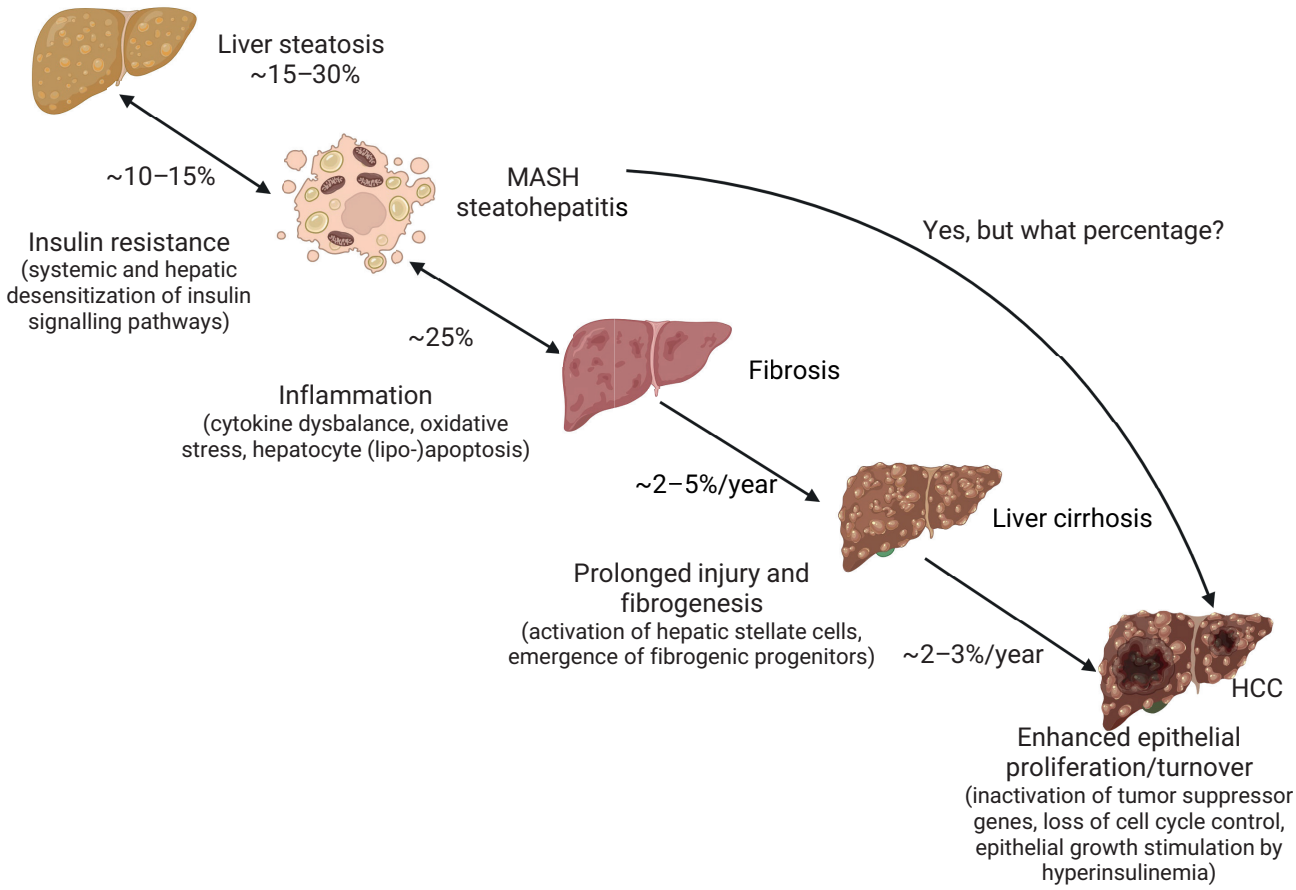


Fig. 3. The natural course of metabolic dysfunction-associated steatohepatitis (MASH)[10]

and HCC.[13] Globally, approx. 20% of NAFLD cases were classified as NASH in 2015, and by 2030, this proportion is projected to increase to 27% according to the Makarov model, reflecting both disease progression and population aging. The incidence of decompensated cirrhosis is projected to increase by 168% to 105,430 cases by 2030, while the incidence of HCC is projected to increase by 137% to 12,240 cases. Deaths from liver disease are projected to increase by 178% to an estimated 78,300 deaths by 2030.[14] Given the increasing morbidity and mortality associated with progressive forms of MASLD, there is growing urgency to identify effective therapies. In March 2024, the U.S. Food and Drug Administration (FDA) approved resmetirom for the treatment of MASH/MASLD, but this drug has not yet been approved in the EU. In routine practice, treatment still relies on management of individual components of metabolic syndrome, including the use of bariatric surgery techniques. However, the long-term effectiveness of these procedures remains unsatisfactory.

SGLT2 inhibitors (gliflozins)

SGLT2 inhibitors (sodium-glucose co-transporter 2 inhibitors (SGLT2i)) are a modern group of drugs originally used as oral antidiabetic agents, either as monotherapy for T2DM or in combination with insulin or other oral antidiabetic agents in the treatment of this disease. Gliflozins block sodium-glucose co-transporter 2 (SGLT2) in the kidney. During blood filtration in the kidney, SGLT2 is responsible for the reabsorption of glucose from urine into the bloodstream. Thus, when SGLT2 activity is blocked, urinary glucose excretion increases and blood glucose concentration decreases. The mechanism of action of gliflozins is independent of insulin.[15]

SGLT2, which is mainly found in the proximal tubule, is responsible for up to 97% of renal glucose reabsorption and approx. 5% of total renal sodium reabsorption. The direct and indirect effects on sodium transport play an important role in cardioprotection and renoprotection; therefore, in addition to diabetes, gliflozins have also found applications in cardiology and nephrology, as shown in Fig. 4.[15,16]

According to the first large completed clinical trial, the EMPA-REG OUTCOME trial published in 2015, compared with placebo, the use of SGLT2i resulted in a significantly lower risk of cardiovascular death, reduced all-cause mortality, and fewer hospitalizations for heart failure.[15] In large cardiovascular trials, beneficial effects on renal function were also observed, which led to studies focusing primarily on renoprotection. The first breakthrough study was the CREDENCE trial, which investigated the effect of canagliflozin on glomerular filtration rate in patients with T2DM. In this study, canagliflozin reduced the risk

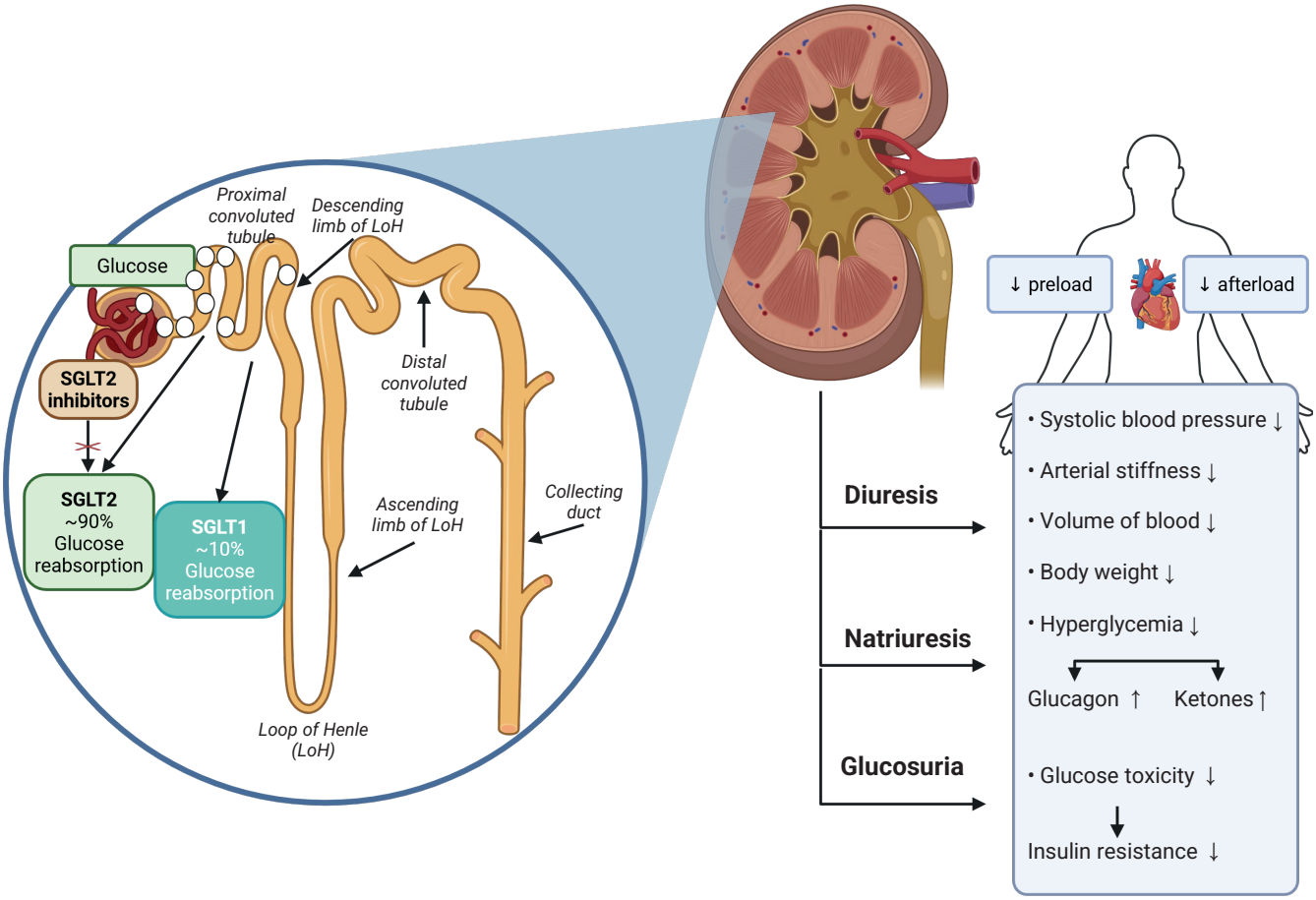


Fig. 4. The mechanism of action of gliflozines[16]

of renal and cardiovascular events independently of baseline glomerular filtration rate.[17]

The beneficial effects of SGLT2i are likely multifactorial and include alterations in arterial stiffness, cardiac function, and cardiac oxygen consumption. In addition to cardiorenal effects, SGLT2i reduce albuminuria and hyperuricemia and favorably affect hyperglycemia, body weight, visceral adiposity, and blood pressure.[15] Other effects of SGLT2i include reduction of oxidative stress and reactive oxygen species (ROS) generated by TNF-α through inhibition of sodium-hydrogen exchange and reduction of intracellular sodium levels.[18]

SGLT2 inhibitors in the treatment of MASLD/MASH

Even in the case of MASLD, a growing number of studies have evaluated the effects of SGLT2i treatment on liver disease. A multicenter retrospective study by Huynh et al. compared 5-year mortality, the composite risk of hepatic decompensation, and the incidence of HCC in patients with diabetes and liver cirrhosis treated with either metformin monotherapy or dual therapy consisting of metformin and SGLT2i. Dual therapy was associated with reduced mortality compared with monotherapy (p = 0.002).[8,19,20]

In a recent meta-analysis evaluating the effects of SGLT2i on liver fibrosis and MASLD, 18 studies published up to December 2023 involving a total of 1,330 patients were included after meeting the eligibility criteria. The meta-analysis concluded that SGLT2i use may lead to modest improvement in hepatic steatosis and/or fibrosis compared with controls in patients with MASLD and T2DM, based on imaging and histopathological biomarkers, with low to moderate certainty of evidence, as shown in Table 2,3.[19,21–26]

Case-control studies suggest that the use of glucagon-like peptide-1 receptor agonists (GLP-1 receptor agonists) or SGLT2i in patients with T2DM may be associated with a reduced risk of hepatic complications, although the only available pilot study of semaglutide in patients with cirrhosis did not demonstrate histological improvement. In the study by Newsome et al., 320 patients (230 of whom had stage F2 or F3 fibrosis) were randomly assigned to receive semaglutide at doses ranging from 0.1 mg to 0.4 mg or placebo. The highest percentage of patients achieving MASH resolution was observed in the 0.4-mg group (59%) compared with the placebo group (17%) (p < 0.001 for semaglutide 0.4 mg vs placebo). Improvement in fibrosis stage occurred in 43% of patients in the 0.4-mg group and in 33% of patients in the placebo group (p = 0.48). Although this trial demonstrated a significantly higher rate

Table 2. Forest plot of comparison: summary of mean difference in controlled attenuation parameter (CAP) post-treatment in patients with non-alcoholic fatty liver disease (NAFLD) randomized to either sodium-glucose co-transporter 2 (SGLT-2) inhibitor or control[19]

Study or subgroup	SGLT-2i mean [dB/m]	SGLT-2i SD [dB/m]	SGLT-2i total	Control mean [dB/m]	Control SD [dB/m]	Control total	Weight	Mean difference (95% CI)	Risk of bias					
									A	B	C	D	E	F
Chehrehgosha et al., 2021[21]	287.8	31.14	35	296.73	40.13	37	21.5%	–8.93 [–25.47, 7.61]	+	+	+	+	+	+
Han et al., 2019[22]	298.6	45.2	29	319.5	37.3	24	11.9%	–20.90 [–43.11, 1.31]	?	+	+	?	+	?
Hu et al., 2020[23]	254.4	31.3	30	264.76	32.74	30	22.3%	–10.36 [–26.57, 5.85]	?	+	+	+	+	?
Shimizu et al., 2019[24]	290.3	72.7	33	311.3	37.3	24	7.0%	–21.00 [–49.95, 7.95]	?	?	+	+	+	?
Taheri et al., 2020[25]	277.7	31.9	43	281.2	34.7	47	31.0%	–3.50 [–17.26, 10.26]	?	?	+	+	+	?
Takeshita et al., 2022[26]	261.8	58.1	20	282.7	38.7	20	6.3%	–20.90 [–51.49, 9.69]	?	+	+	+	+	?
Total (95 % CI)			190			182	100%	–10.59 [–18.25, –2.92]						

Heterogeneity: $Tau^2 = 0.00$; $\chi^2 = 2.82$, df = 5 (P = 0.73); $I^2 = 0\%$; overall effect: Z = 2.71 (p = 0.007); df – degrees of freedom; SD – standard deviation; 95% CI – 95% confidence interval.
Test for subgroup differences: Not applicable.
Risk of bias legend:
A – Bias arising from randomization process
B – Bias due to deviations from intended interventions
C – Bias due to missing outcome data
D – Bias in measurement of the outcome
E – Bias in the selection of the reported result
F – Overall bias

Table 3. Forest plot of comparison (sensitivity analysis): summary of mean difference in controlled attenuation parameter (CAP) post-treatment in patients with non-alcoholic fatty liver disease (NAFLD) randomized to either SGLT-2 inhibitor or control[19]

Study or subgroup	SGLT-2i mean [dB/m]	SGLT-2i SD [dB/m]	SGLT-2i total	Control mean [dB/m]	Control SD [dB/m]	Control total	Weight	Mean difference (95% CI)	Risk of bias					
									A	B	C	D	E	F
Chehrehgosha et al., 2021[21]	287.8	31.14	35	296.73	40.13	37	31.1%	–8.93 [–25.47, 7.61]	+	+	+	+	+	+
Han et al., 2019[22]	298.6	45.2	29	319.5	37.3	24	17.3%	–20.90 [–43.11, 1.31]	?	+	+	?	+	?
Hu et al., 2020[23]	254.4	31.3	30	264.76	32.74	30	32.4%	–10.36 [–26.57, 5.85]	?	+	+	+	+	?
Shimizu et al., 2019[24]	290.3	72.7	33	311.3	37.3	24	10.2%	–21.00 [–49.95, 7.95]	?	?	+	+	+	?
Taheri et al., 2020 excluded[25]	277.7	31.9	43	281.2	34.7	47	31.0%	–3.50 [–17.26, 10.26]	?	?	+	+	+	?
Takeshita et al., 2022[26]	261.8	58.1	20	282.7	38.7	20	9.1%	–20.90 [–51.49, 9.69]	?	+	+	+	+	?
Total (95 % CI)			147			135	100%	-13.77 [-23.00, -4.55]						

Heterogeneity: $Tau^2 = 0.00$; $\chi^2 = 1.34$, df = 5 (p = 0.85); $I^2 = 0\%$; overall effect: Z = 2.93 (p = 0.003); df – degrees of freedom; SD – standard deviation; 95% CI – 95% confidence interval.
Test for subgroup differences: Not applicable.
Risk of bias legend:
A – Bias arising from randomization process
B – Bias due to deviations from intended interventions
C – Bias due to missing outcome data
D – Bias in measurement of the outcome
E – Bias in the selection of the reported result
F – Overall bias

of MASH resolution with semaglutide compared with placebo, it did not show a significant improvement in fibrosis stage in patients with MASH.[27,28]

There is currently an ongoing multicenter, randomized, double-blind, placebo-controlled phase 3 trial involving patients with MASH and moderate or advanced liver fibrosis who once weekly receive semaglutide at a dose of 2.4 mg or placebo. Interim trial results have demonstrated statistically significant improvement in liver histological findings. However, further research and additional data are required to confirm these results.[29]

In a meta-analysis by Zhou et al., a total of 686 patients from 8 studies were included. Compared with the control group, SGLT2i significantly reduced controlled attenuation parameter (CAP) values (mean difference (MD) = –12.80, 95% confidence interval (95% CI): [–20.57, –5.03], p = 0.001; Fig. 5) and liver stiffness measurement (LSM) values (MD = –0.82, 95% CI: [–1.38, –0.25], p = 0.005).[30]

According to the 2024 European guidelines for MASLD, there is currently insufficient evidence to recommend SGLT2i or metformin as specific treatments for MASH. However, their use in patients with MASLD is considered safe within currently approved indications. The guidelines indicate that SGLT2i are safe in patients with Child–Pugh class A and B cirrhosis and may be used according to their approved indications.[21] Sufficient evidence

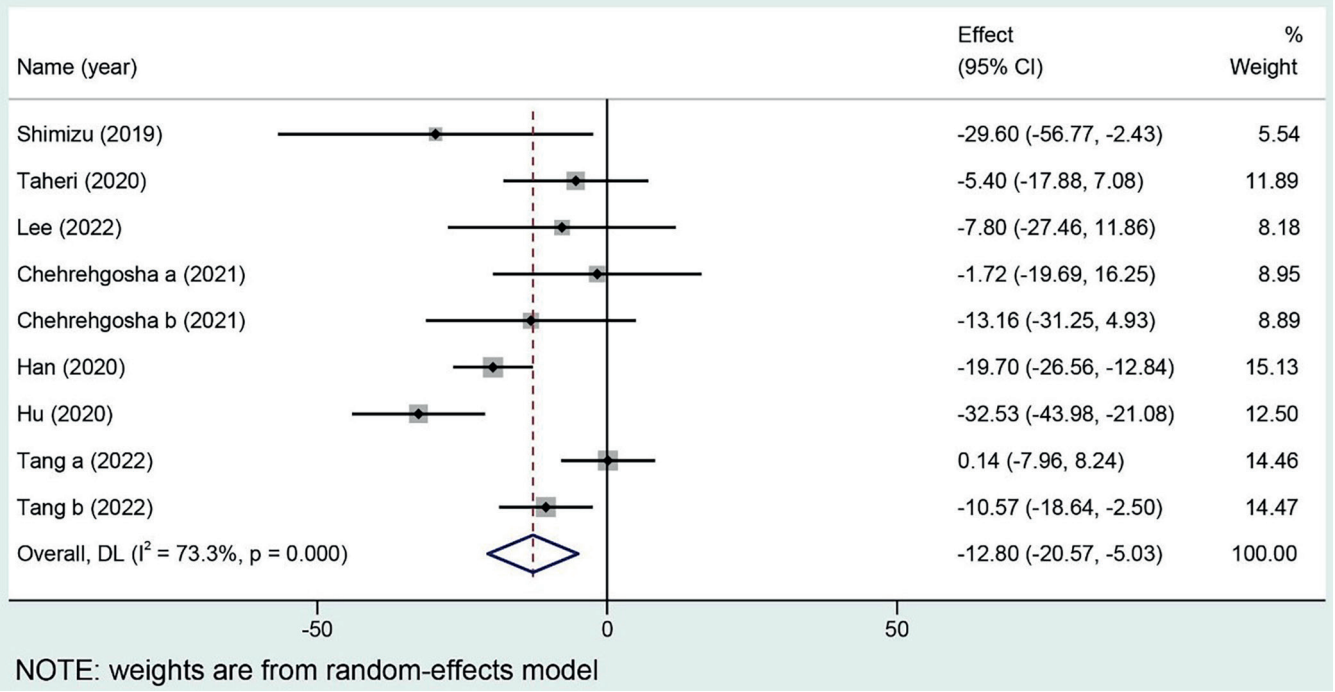


Fig. 5. Forest plot for the overall effects on the treatment of controlled attenuation parameter (CAP)[30]

supporting the effects of SGLT2i on reducing inflammatory activity and fibrosis in patients with histologically confirmed MASH is still lacking.

Limitations of the study

The main limitation of this review is that it falls within the category of narrative or scoping reviews and provides only a brief overview of the topic. It is not a protocol-driven, in-depth data analysis and may therefore be subject to limitations such as bias.

Conclusions

The importance of SGLT2i in diabetes, cardiology, and nephrology has already been demonstrated, and their use is becoming increasingly common across various clinical indications; however, their role in the treatment of MASLD/MASH has not yet been fully clarified. Nevertheless, SGLT2i are likely to be encountered in patients with MASLD/MASH during treatment of associated diseases related to metabolic syndrome and its complications. Some evidence suggests that SGLT2i may have beneficial effects on the course of MASLD/MASH, including potential effects on fibrosis, steatosis, and oxidative stress, at least through indirect mechanisms. However, sufficient evidence evaluating the effects of SGLT2i on reducing histologically confirmed inflammatory activity and liver fibrosis in MASH is still lacking. Further research focusing on histological assessment of the effects of SGLT2i is needed.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Jan Paleček https://orcid.org/0009-0004-0302-8786
Petr Hříbek https://orcid.org/0000-0002-8090-4389
Denisa Janíčková Ždarská https://orcid.org/0000-0002-8143-7276
Marie Nováková https://orcid.org/0009-0008-0692-8957
Pavel Skořepa https://orcid.org/0000-0002-9401-1848
Petr Urbánek https://orcid.org/0000-0002-1506-1135

References

1. Younossi ZM, Golabi P, Paik JM, Henry A, Van Dongen C, Henry L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): A systematic review. *Hepatology*. 2023;77(4):1335–1347. doi:10.1097/HEP.0000000000000004
2. Kanwal F, Neuschwander-Tetri BA, Loomba R, Rinella ME. Metabolic dysfunction-associated steatotic liver disease: Update and impact of new nomenclature on the American Association for the Study of Liver Diseases practice guidance on nonalcoholic fatty liver disease. *Hepatology*. 2024;79(5):1212–1219. doi:10.1097/HEP.0000000000000670
3. Gofton C, Upendran Y, Zheng MH, George J. MAFLD: How is it different from NAFLD? *Clin Mol Hepatol*. 2023;29(Suppl):S17–S31. doi:10.3350/cmh.2022.0367
4. Eslam M, Newsome PN, Sarin SK, et al. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *J Hepatol*. 2020;73(1):202–209. doi:10.1016/j.jhep.2020.03.039
5. Vallet-Pichard A, Mallet V, Nalpas B, et al. FIB-4: An inexpensive and accurate marker of fibrosis in HCV infection. Comparison with liver biopsy and fibrotest. *Hepatology*. 2007;46(1):32–36. doi:10.1002/hep.21669
6. Šmíd V, Dvořák K. Doplnění Doporučeného postupu České hepatologické společnosti ČLS JEP pro diagnostiku a léčbu NAFLD: FIB-4 index – interpretace získaného výsledku a doporučení dalšího postupu. Prague, Czech Republic: Czech Hepatological Society; 2023. https://www.ces-hep.cz/d-doc/1/251/638804887935200000/doplneni-dop-postupu-fib-4-schvaleno-05-12-2023.pdf. Accessed August 13, 2024.

7. Brůha R, Dvořák K, Fejfar T, Šmíd V, Trunečka P. Doporučený postup České hepatologické společnosti ČLS JEP pro diagnostiku a léčbu Nealkoholové tukové choroby jater (NAFLD). Prague, Czech Republic: Czech Hepatological Society; 2023. https://www.ces-hep.cz/d-doc/1/239/638804887886233333/nafld-dop-postup-chs.pdf. Accessed August 13, 2024.
8. Lu S, Wang Y, Liu J. Tumor necrosis factor-α signaling in nonalcoholic steatohepatitis and targeted therapies. *J Genet Genomics*. 2022;49(4):269–278. doi:10.1016/j.jgg.2021.09.009
9. Li X, Wang H. Multiple organs involved in the pathogenesis of nonalcoholic fatty liver disease. *Cell Biosci*. 2020;10(1):140. doi:10.1186/s13578-020-00507-y
10. Schuppan D, Schattenberg JM. Non-alcoholic steatohepatitis: Pathogenesis and novel therapeutic approaches. *J Gastroenterol Hepatol*. 2013;28(Suppl 1):68–76. doi:10.1111/jgh.12212
11. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
12. Simon TG, Roelstraete B, Hagström H, Loomba R, Ludvigsson JF. Progression of non-alcoholic fatty liver disease and long-term outcomes: A nationwide paired liver biopsy cohort study. *J Hepatol*. 2023;79(6):1366–1373. doi:10.1016/j.jhep.2023.08.008
13. Golabi P, Paik JM, Al-Qahtani S, Younossi Y, Tuncer G, Younossi ZM. Burden of non-alcoholic fatty liver disease in Asia, the Middle East and North Africa: Data from Global Burden of Disease 2009–2019. *J Hepatol*. 2021;75(4):795–809. doi:10.1016/j.jhep.2021.05.022
14. Estes C, Razavi H, Loomba R, Younossi Z, Sanyal AJ. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology*. 2018;67(1):123–133. doi:10.1002/hep.29466
15. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. *N Engl J Med*. 2015;373(22):2117–2128. doi:10.1056/NEJMoa1504720
16. Dutka M, Bobiński R, Ulman-Włodarz I, et al. Sodium glucose cotransporter 2 inhibitors: Mechanisms of action in heart failure. *Heart Fail Rev*. 2021;26(3):603–622. doi:10.1007/s10741-020-10041-1
17. Jardine MJ, Zhou Z, Mahaffey KW, et al. Renal, cardiovascular, and safety outcomes of canagliflozin by baseline kidney function: A secondary analysis of the CREDENCE randomized trial. *J Am Soc Nephrol*. 2020;31(5):1128–1139. doi:10.1681/ASN.2019111168
18. Uthman L, Li X, Baartscheer A, et al. Empagliflozin reduces oxidative stress through inhibition of the novel inflammation/NHE/[Na+]c/ROS-pathway in human endothelial cells. *Biomed Pharmacother*. 2022;146:112515. doi:10.1016/j.biopha.2021.112515
19. Ong Lopez AMC, Pajimna JAT. Efficacy of sodium glucose cotransporter 2 inhibitors on hepatic fibrosis and steatosis in non-alcoholic fatty liver disease: An updated systematic review and meta-analysis. *Sci Rep*. 2024;14(1):2122. doi:10.1038/s41598-024-52603-5
20. Huynh DJ, Renelus BD, Jamorabo DS. Reduced mortality and morbidity associated with metformin and SGLT2 inhibitor therapy in patients with type 2 diabetes mellitus and cirrhosis. *BMC Gastroenterol*. 2023;23(1):450. doi:10.1186/s12876-023-03085-8
21. Chehrehgosha H, Sohrabi MR, Ismail-Beigi F, et al. Empagliflozin improves liver steatosis and fibrosis in patients with non-alcoholic fatty liver disease and type 2 diabetes: A randomized, double-blind, placebo-controlled clinical trial. *Diabetes Ther*. 2021;12(3):843–861. doi:10.1007/s13300-021-01011-3
22. Han E, Lee Y ho, Lee BW, Kang ES, Cha BS. Ipragliflozin additively ameliorates non-alcoholic fatty liver disease in patients with type 2 diabetes controlled with metformin and pioglitazone: A 24-week randomized controlled trial. *J Clin Med*. 2020;9(1):259. doi:10.3390/jcm9010259
23. Hu CL, Wang YC, Xi Y, Yao XM. Dapagliflozin therapy curative effect observation on nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus. *Indian J Pharm Sci*. 2020;Special Issue 7:122–129. https://www.ijpsonline.com/articles/dapagliflozin-therapy-curative-effect-observation-onnonalcoholic-fatty-liver-disease-in-patients-with-type-2diabetes-mel.pdf.
24. Shimizu M, Suzuki K, Kato K, et al. Evaluation of the effects of dapagliflozin, a sodium-glucose co-transporter-2 inhibitor, on hepatic steatosis and fibrosis using transient elastography in patients with type 2 diabetes and non-alcoholic fatty liver disease. *Diabetes Obes Metab*. 2019;21(2):285–292. doi:10.1111/dom.13520
25. Taheri H, Malek M, Ismail-Beigi F, et al. Effect of empagliflozin on liver steatosis and fibrosis in patients with non-alcoholic fatty liver disease without diabetes: A randomized, double-blind, placebo-controlled trial. *Adv Ther*. 2020;37(11):4697–4708. doi:10.1007/s12325-020-01498-5
26. Takeshita Y, Honda M, Harada K, et al. Comparison of tofogliflozin and glimepiride effects on nonalcoholic fatty liver disease in participants with type 2 diabetes: A randomized, 48-week, open-label, active-controlled trial. *Diabetes Care*. 2022;45(9):2064–2075. doi:10.2337/dc21-2049
27. European Association for the Study of the Liver (EASL), European Association for the Study of Diabetes (EASD), European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines on the Management of Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD). *Obes Facts*. 2024;17(4):374–444. doi:10.1159/000539371
28. Newsome PN, Buchholtz K, Cusi K, et al. A placebo-controlled trial of subcutaneous semaglutide in nonalcoholic steatohepatitis. *N Engl J Med*. 2021;384(12):1113–1124. doi:10.1056/NEJMoa2028395
29. Sanyal AJ, Newsome PN, Kliers I, et al. Phase 3 trial of semaglutide in metabolic dysfunction-associated steatohepatitis. *N Engl J Med*. 2025;392(21):2089–2099. doi:10.1056/NEJMoa2413258
30. Zhou P, Tan Y, Hao Z, Xu W, Zhou X, Yu J. Effects of SGLT2 inhibitors on hepatic fibrosis and steatosis: A systematic review and meta-analysis. *Front Endocrinol (Lausanne)*. 2023;14:1144838. doi:10.3389/fendo.2023.1144838

Formaldehyde neurotoxicity: Effects on the mammalian brain, cognitive function, and neurodegenerative risk. A scoping review

Mateusz Drążyk[1,A–D,F], Zuzanna Pyc[2,B–D], Szymon J. Pietrzyk[1,B–D], Antonina Gajda-Janiak[1,B–D], Filip Godziszewski[1,B–D], Oliwier Pioterek[1,B,E,F], Michał Tulski[1,B,E], Mateusz Mazurek[1,A,E,F], Zygmunt A. Domagała[3,A,E,F]

[1] Student Scientific Association of Clinical and Dissecting Anatomy, Wroclaw Medical University, Poland
[2] Student Scientific Association of Biomedical and Environmental Analyses, Wroclaw Medical University, Poland
[3] Division of Anatomy, Wroclaw Medical University, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1093–1107

Address for correspondence
Mateusz Mazurek
E-mail: mateusz.mazurek@student.umw.edu.pl

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
We would like to thank the technical personnel of our department, especially Mr. Mirosław Łukaszun and Mr. Zbigniew Staszewski. The idea for this review is rooted in their everyday work with us in the anatomy department.

Received on March 25, 2025
Reviewed on July 25, 2025
Accepted on August 17, 2025

Published online on June 25, 2026

Abstract

Aqueous formaldehyde (FA) solution, known as formalin, is currently the primary agent used for preserving tissue samples and anatomical specimens. Formaldehyde is widely used in laboratories and the chemical industry; it also occurs as an air pollutant and endogenous cellular metabolite. The potential carcinogenic effects of formalin on the respiratory tract are well documented. A less recognized consequence of occupational exposure to FA is its detrimental effect on the central nervous system (CNS) and brain function. A literature review was conducted to investigate the effects of FA on the brain. Five databases were searched: PubMed, Web of Science (WoS), Embase, ScienceDirect, and Google Scholar. To describe the effects of FA exposure and endogenous FA generation, 35 relevant publications were collected and analyzed. The literature review demonstrated that inhalation is the most common route of FA exposure. Several studies have shown that FA may cause hippocampal damage, disrupt melatonin secretion, and induce a wide range of cognitive disorders with varying characteristics and severity. These disorders include memory impairment, disturbances in balance and spatial orientation, learning difficulties, sleep disturbances, impaired judgment, and prolonged reaction times to stimuli. Increased endogenous FA concentration has also been associated with a higher risk of neurodegenerative diseases, such as Alzheimer's disease (AD) and amyotrophic lateral sclerosis. The literature analysis demonstrated the high neurotoxicity of FA, which may lead to numerous neuropsychiatric disorders. We aim to draw attention to the risks associated with the routine use of formalin, particularly among anatomists and pathologists, and to encourage consideration of less harmful alternative preservation agents.

Key words: formaldehyde, occupational exposure, cognitive impairment, neurotoxicity, neurodegenerative diseases

Cite as
Drążyk M, Pyc Z, Pietrzyk SJ, et al. Formaldehyde neurotoxicity: Effects on the mammalian brain, cognitive function, and neurodegenerative risk. A scoping review. *Adv Clin Exp Med*. 2026;35(6):1093–1107. doi:10.17219/acem/209617

DOI
10.17219/acem/209617

Highlights

- Formaldehyde exposure may contribute to deterioration in mental health.
- Research suggests that formaldehyde exposure may affect mood, cognitive function, and sleep quality.
- Formaldehyde contributes to oxidative stress and apoptosis in neural cells, potentially leading to cognitive decline.
- Reducing or eliminating formaldehyde use in research, particularly in medical settings, should be a priority for occupational safety and health.

Introduction

Formaldehyde (FA) has been known for more than 150 years. It was first synthesized in gaseous form in 1859 by Alexandr Butlerov. Nearly a decade later, in 1868, German scientist August Wilhelm von Hofmann described its chemical structure.[1] Since then, numerous applications of this substance have been identified. Formaldehyde is commonly used in the textile industry and carpentry as a substrate for the production of resins and glues and, in the form of an aqueous solution known as formalin, is widely recognized as the main ingredient of body-preserving solutions used for anatomical and funeral purposes. The use of FA in anatomy began at the end of the 19th century,[2] and since then it has continued to be used as a fixative for tissue preservation by pathologists and histologists.[3]

Formaldehyde is also present in the air we breathe as an environmental pollutant,[4,5] and is a component of cigarette smoke and vehicle exhaust fumes.[6–8] It is widely used in industrial production as a substrate in the chemical industry[9] and can also be found in cosmetic products.[10,11] A lesser-known fact is that FA is one of the products of biochemical pathways in the human body.[12] These examples demonstrate that the human population is constantly exposed to FA. The level of exposure may vary depending on occupation or place of residence, but exposure is unavoidable.

In 2004, FA was classified as a carcinogen by the International Agency for Research on Cancer (IARC).[13] There is sufficient evidence suggesting that it increases the risk of nasopharyngeal cancer and leukemia.[14] Several studies have also investigated whether FA exposure may increase the risk of brain tumor development.[15] Formaldehyde has been identified as a neurotoxic substance in numerous in vivo and ex vivo studies. Although the carcinogenic properties of FA are widely recognized, it continues to be extensively used. Its low cost and ease of use contribute to its popularity in pathology, histology, and education (e.g., gross anatomy courses). In contrast to other fixatives, FA penetrates tissue easily, acts rapidly, and does not require unusual storage conditions. However, relatively few studies have examined the neurotoxicity of FA and its potential effects on cognition and behavior.

Objectives

In this study, we examined the relationship between neuronal damage and death and reduced brain function associated with exposure to FA and endogenous FA production. In particular, we focused on cognitive impairment, including FA-induced alterations in memory and attention span. We also investigated whether FA may contribute to the development of mood disorders.

Furthermore, we investigated the molecular mechanisms of FA action, its biological and psychological implications, and its potential contribution to the development of neurodegenerative diseases. Considering the contribution of FA to carcinogenesis, we hypothesized that FA may also adversely affect neuronal health and neuropsychological functioning. Our investigation aimed to provide an objective and comprehensive evaluation of the available scientific evidence regarding FA and nervous system health.

Materials and methods

Eligibility criteria

To explore the influence of FA on the central nervous system (CNS), we reviewed the available literature. Only original full-text articles written in English were included. Review articles were excluded. No restrictions regarding publication date were applied. We included studies conducted in humans, animal models, and cell cultures. We searched for information regarding changes in cognitive function, cognitive decline, and behavioral alterations following FA exposure. Additional selected articles described FA metabolism and biochemical pathways in organisms, particularly mechanisms underlying neuronal damage and alterations in neuronal physiology. Finally, we also included articles addressing the influence of exogenous and endogenous FA on neurodegenerative diseases.

Information sources and search strategy

Our search was conducted in 5 databases: PubMed, Web of Science (WoS), Embase, ScienceDirect, and Google Scholar. The search began on November 27, 2023, and ended on December 7, 2023.

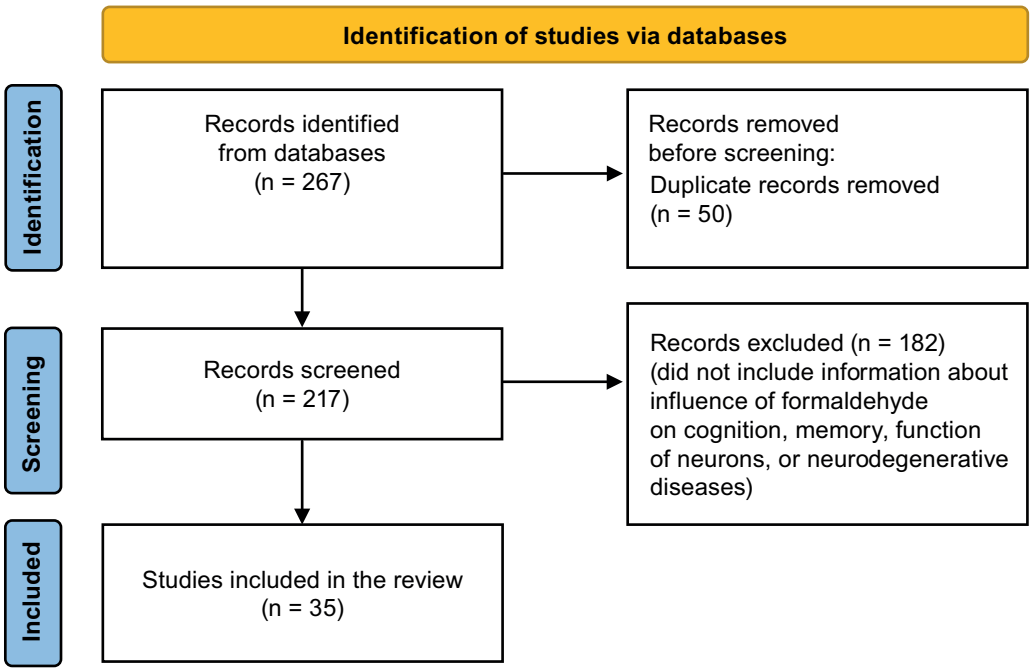


Fig. 1. Flow diagram depicting the study selection process

We defined search phrases that were subsequently used to identify articles describing FA and the CNS. The search phrases consisted of the constant term "formaldehyde exposure", combined with the following additional terms: "human", "brain damage", "cognition", "cognitive functions", "dementia", "intelligence", "memory", "memory loss", "neurobehavioral impairment", "neurological disorders", "neuropsychiatric disorders", "neurotoxicity", and "induced neuron apoptosis".

Data collection and outcomes

In the initial search, based on title screening conducted in the above-mentioned databases, we identified 267 articles. After removing duplicates, 217 papers remained. Subsequently, abstracts of these articles were screened, which reduced the number of eligible studies to 35. Studies conducted on animal models comprised 18 of the 35 articles included in the final pool of selected studies (Fig. 1).

In the next step, data were extracted from the selected articles, which were subsequently reviewed and summarized. Using an extraction form created in Microsoft Word 365 (Microsoft Corp., Redmond, USA), we collected all data related to the above-mentioned terms. We searched for information regarding the effects of FA on neuronal function and the CNS. The collected data also addressed the neurotoxicity of FA. We then examined the impact of neuronal damage or death on cognitive functions, including memory, learning processes, spatial orientation, reaction time, abstract thinking, and judgment, as well as prolonged reaction times to stimuli. We were also interested in attention span, behavioral changes, and affective disorders following exposure. Although the focus was mainly on acute and chronic FA exposure, we also considered endogenous FA formation in the context of cognitive decline and neurodegenerative diseases.

The collected data were subsequently organized into 4 main subject areas: neurotoxicity, the influence of FA on memory, the influence of FA on other cognitive functions, and the development of FA-related neurodegenerative diseases. The summarized data from all studies included in this review are presented in Table 1,[16–24] Table 2,[24–43] and Table 3.[4,5,44–52]

Results and discussion

Neurotoxicity of formaldehyde

Formaldehyde readily crosses the blood–brain barrier (BBB).[53] Once in the brain, it may damage neurons. For example, the methyl donor S-adenosylmethionine (SAM) in PC12 cell cultures derived from patients with Parkinson's disease induces a cascade of methanol-derived reactions leading to the formation of endogenous FA. Formaldehyde, produced through oxidation of methanol during SAM-induced reactions, showed marked toxicity toward neuron-like PC12 model cells, which are catecholaminergic pheochromocytoma cells. Under the same conditions, methanol and formic acid did not exhibit comparable toxicity. Furthermore, FA induced a considerably weaker toxic effect in C6 glial cell cultures – mainly excessive protein carboxymethylation and inhibition of tyrosine hydroxylase expression – indicating that neurons may be more susceptible to FA toxicity than glial cells. However, prolonged incubation resulted in decreased viability of both PC12 and C6 cells, suggesting that exposure duration plays a crucial role in the toxic effects of FA on human neuronal and glial cells.[19]

Table 1. Formaldehyde-induced mechanisms of neurotoxicity

Reference	Model (human/ animal/cell line)	Size – control group	Size – experimental group(s)	Type and duration of exposure	Mechanism of neurotoxicity	Results
Askar and Halloull (2018)[16]	rat	10 – group 1	10 (group 2), 10 (group 3)	Airborne FA exposure; Group 2: FA exposure for 6 h/day, 5 days/week for 12 weeks, Group 3: FA exposure for 6 h/day 5 days/week + omega-3-wheat germ oil supplement, for 12 weeks	Increased production of ROS; upregulation of BAX	Damage to neurons; Intensified OS, apoptosis
Ciftci et al. (2015)[17]	rat	10	10 (group 2), 10 (group 3)	Group 2: 37% FA injection, 9 mg/kg for 15 days; Group 3: 37% FA injection, 9 mg/kg paired with 100 mg/kg curcumin every day, for 15 days	Increased levels of 8-OHDG in the brain tissue and urine of rats from group 2	Oxidative stress in rats' brains
Gurel et al. (2005)[18]	rat	13	18 – FA, 18 – FA + vitamin E	Intraperitoneal injection 10 mg/kg of FA for 10 days	Lipid and protein peroxidation; formation of ROS; Decreased SOD and cat activity in brain tissue	Increased OS; Degenerative changes in neurons
Lee et al. (2008)[19]	PC12 and C6 cell lines	untreated cells	cells treated with FA, methanol and formic acid	Incubation of cells in increasing concentrations of FA for 24 h	Excessive protein carboxymethylation; Inhibition of tyrosine hydroxylase expression	Death of neurons
Li et al. (2020)[20]	rat	11	11 FA-treated (group 1) 11 FA-NHS-treated (group 2)	Intracerebroventricular injection 10 µmol/day FA for 7 days	Significant increase in MDA level and decrease in SOS activity in hippocampal cells; Reduction of expression of BDNF protein; Significant increase in number of apoptotic neurons	Promotion of OS and apoptosis in hippocampal neurons
Li et al. (2021)[21]	mouse HT22 cell line	untreated cells	cells treated with FA, concentration of 0.1, 0.5, or 1.0 mm	Cells cultured with FA for 6 h	Induction of changes leading to ferroptosis; Upregulation of Warburg effect	Reduced cell viability and increased mortality; Iron-depended apoptosis in neurons; Increase in oxidative stress
Tang et al. (2013)[22]	PC12 cell line	untreated cells	cells treated with FA	FA present in the phosphate-buffered saline (PBS), 5–15 min incubation	Decrease in anti-apoptotic BCL2 protein; Release of CYT-C leading to apoptosis; Increase in ROS levels; concentration dependent inhibition of paroxonase-1 expression and activity	Induction of apoptosis in PC12 cells; Increase in oxidative stress
Tang et al. (2013)[22]	PC12 cell line	untreated cells	cells treated with FA	120 and 240 mmol/L FA concentrations, incubation for 24 h	Significantly decreased NO and H2S production and CBS expression in cells	Reduced cells' viability; neurotoxicity of FA; Increased ROS accumulation in treated cells
Umansky et al. (2022)[23]	NALM6 and HCT116 cell line	untreated cells	cells treated with FA	Incubation of cell cultures in FA	FA blocks GSH antioxidant capability	Impairment of anti-oxidant defense; Accumulation of ROS
Zendehdel et al. (2016)[24]	human	32	35	Occupational exposure to FA vapors, exposure predates beginning of the study	Higher levels of ache activity in erythrocytes of experimental group	Possible connection with cognitive decline

FA – formaldehyde; ROS – reactive oxygen species; OS – oxidative stress; 8-OHdG – 8-hydroxy-2′-deoxyguanosine; SOD – superoxide dismutase; CAT – catalase; PC12 – pheochromocytoma cell line; C6 – rat glial tumor cell line; MDA – malondialdehyde; BDNF – brain-derived neurotrophic factor; HT22 – mouse hippocampal neuronal cell line; Cyt-C – cytochrome C; NO – nitric oxide; H_2S – hydrogen sulfide; CBS – cystathionine β-synthase; GSH – glutathione; AChE – acetylcholinesterase; NHS – N-hydroxysuccinimide.

Table 2. Influence of FA on cognitive functions

Reference	Model (human/ animal/cell line)	Size (control group)	Size – experimental group(s)	Type and duration of exposure	Results
Bach et al. (1990)[25]	human	not disclosed	61: 32 – workers occupationally exposed to FA 29 – unexposed occupationally	FA concentrations: 0.0, 0.15, 0.4, or 1.2 mg/m^3 climate chamber for 5.5 h for 16 days	possible distractive sensory irritation due to FA exposure
Dou et al. (2012)[26]	mouse	12	12 – 1 mg/m^3 FA concentration, 12 – 3 mg/m^3 FA concentration, 12 – 5 mg/m^3 FA concentration	airborne FA exposure	impaired memory and learning ability in mice exposed to FA, increasing with FA concentration
Kilburn et al. (1987)[27]	human	not disclosed	305	airborne FA exposure, exposure predated beginning of the study, total duration of exposure varying between individuals	reduced performance on story memory, visual memory, digit span, and sharpened-Romberg
Kilburn et al. (1983)[28]	human	not disclosed	420 – FA concentration ranging from 0.4 ppm to 5 ppm	airborne FA exposure, exposure predates beginning of the study, total duration of exposure varying between individuals	impaired short-term memory impaired ability to concentrate
Kilburn et al. (1985)[29]	human	26	45 – exposed to FA during fiberglass batt-making, 18 – exposed to FA during fixation of the tissue in histology laboratories	airborne FA exposure, exposure predates beginning of the study, total duration of exposure varying between individuals	impaired memory, loss of concentration, sleep disturbances, dizziness, mood swings
Kilburn et al. (1985)[30]	human	56 – only women	76 – only women	airborne FA exposure, exposure predated beginning of the study, total duration of exposure varying between individuals	impaired memory, mainly short-term; loss of concentration, sleep disturbances, mood swings, dizziness and headaches, loss of balance
Letellier et al. (2022)[31]	human	~69,296	~6,026	occupational exposure, exposure predates the beginning of the study, total duration of exposure varying between individuals	higher risk of global cognitive impairment in exposed population; the severity of cognitive impairment increasing in correlation with level of exposure to FA
Li et al. (2020)[32]	mouse	10	10–0.5 mg/m^3 FA concentration, 10–3.0 mg/m^3 FA concentration	airborne FA exposure	impaired memory and learning ability in mice exposed to FA
Li et al. (2016)[33]	mouse	15	15–1 ppm FA concentration, 15–2 ppm FA concentration	airborne FA exposure, 2 h a day for 7 days	impaired memory and novel objects recognition in mice exposed to higher (5 ppm) doses of FA, depression- and anxiety-like behavior
Lu et al. (2008)[34]	mouse	5	5–1 mg/m^3 FA concentration, 5–3 mg/m^3 FA concentration	airborne FA exposure, 6 h a day for 7 days	impaired memory, impaired ability to learn
Malek et al. (2003)[35]	rat	not disclosed, total number of rats in both groups was 120	not disclosed, total number of rats in both groups was 120, animals in experimental group were divided to smaller assemblies with different levels of FA exposure	airborne FA ingestion, 10 h a day for 7 days	impaired memory, impaired ability to learn
Marceaux et al. (2008)[36]	human	15	5	airborne FA ingestion through smoking marihuana cigarettes, laced with PCP; exposure predates the beginning of the study, total duration of exposure varying between individuals	vocabulary decline, impairment of abstract thinking

Table 2. Influence of formaldehyde on cognitive functions – cont.

Reference	Model (human/ animal/cell line)	Size (control group)	Size – experimental group(s)	Type and duration of exposure	Results
Mei et al. (2016)[37]	mouse	8	3 groups of n = 8 group 2 – gaseous FA 3.0 mg/m^3, group 3 – melatonin injection, group 4 – FA + melatonin	airborne FA ingestion, 8 h for 7 days (groups 2 and 4)	lower spatial learning ability, impaired memory retrieval ability
Perna et al. (2001) [38]	human	not disclosed	1	acute occupational exposure	impaired general memory, impaired ability to acquire new information
Pitten et al. (2000)[39]	rat	14	13–2.6 ppm FA concentration, 13–4.6 ppm FA concentration	airborne FA ingestion, 10 min a day for 90 days	increasing with level of FA exposure: impaired memory formation, impaired learning ability
Schenker et al. (1982)[40]	human	not disclosed	24	airborne FA ingestion, exposure predates beginning of the study, exposure duration ranging from 6 months up to 3 years	impaired memory, impaired attention span
Tang et al. (2013)[41]	rat	not disclosed	not disclosed; FA concentration: 0.1, 1, or 10 µmol/2.5 µL	FA injection, 7 days	impaired ability to learn
Tong et al. (2013)[42]	rat, rat hippocampal neurons and astrocytes culture, human brain samples (healthy and from those who suffered from AD)	not disclosed not applicable not applicable	not disclosed not applicable not applicable	rats: intraperitoneally injected solution 60 mg/kg acute for 7 days chronic for 30 days	suppression of long-term potentiation; NMDA receptor inhibition, resulting in memory loss or memory formation disturbance
Tong et al. (2013)[43]	rat, human SY5Y cell culture, rat hippocampal neuron cell line	not disclosed not applicable not applicable	not disclosed not applicable not applicable	FA injection	impaired memory in rats exposed to FA, decreased methylation, re-methylation of the DNA in cells from cell cultures
Zendehdel, et al. (2016)[24]	human	32	35	airborne FA ingestion during preparation of melanin tableware, exposure predates the beginning of the study, total duration of exposure varying between individuals	concentration of active AChE particles in the blood increased

FA – formaldehyde; ROS – reactive oxygen species; OS – oxidative stress; 8-OHdG – 8-hydroxy-2′-deoxyguanosine; SOD – superoxide dismutase; CAT – catalase; PC12 – pheochromocytoma cell line; C6 – rat glial tumor cell line; MDA – malondialdehyde; BDNF – brain-derived neurotrophic factor; HT22 – mouse hippocampal neuronal cell line; Cyt-C – cytochrome C; NO – nitric oxide; H_2S – hydrogen sulfide; CBS – cystathionine β-synthase; GSH – glutathione; AChE – acetylcholinesterase; NHS – N-hydroxysuccinimide; NMDA – N-methyl-D-aspartate; SH-SY5Y – human neuroblastoma cell line; AD – Alzheimer's disease.

Moreover, numerous studies conducted in laboratory animals have demonstrated visible alterations in brain neuronal tissue following FA exposure. Detected abnormalities included increased reactive oxygen species (ROS) formation,[16,18] other manifestations of oxidative stress (OS),[17] neuronal apoptosis,[16,20] and neuronal degeneration (Fig. 2,3).[18]

Oxidative damage induced by FA

Following FA exposure, numerous biochemical changes indicative of redox imbalance have been observed across different studies. Levels of malondialdehyde (MDA) in the frontal cortex and hippocampus increased after FA exposure.[16,18,21] Malondialdehyde is a product of lipid peroxidation and is considered one of the most important markers of OS.[54] Increased MDA levels have also been observed under controlled conditions simulating occupational exposure.[16] In the frontal cortex and hippocampus, catalase and superoxide dismutase (SOD) activity decreased following FA exposure.[16,18] Both enzymes are involved in defense against OS – catalase through decomposition of hydrogen peroxide (H_2O_2)[55] and SOD through dismutation of superoxide anion radicals into H_2O_2.[56] Rats exposed to FA also exhibited elevated

Table 3. Influence of FA on the development of neurodegenerative diseases

Reference	Model (human/ animal/cell line)	Size – control group	Size – experimental group(s)	Type and duration of exposure	Symptoms, results of study
Calderon-Garciduenas et al. (2015)[4]	human	44 – from group 2	95 67 (group 2) 28 (group 1)	airborne FA ingestion correlated with air pollution, exposure predates beginning of the study	brain imbalance impacting levels of oxidative stress, immune response and inflammation in neural tissue; higher prevalence of tight junction antibodies and neural antibodies; autoimmune response and higher risk of neuroinflammation, development of AD and Parkinson's disease
Hanna et al. (1999)[44]	human	not included in the study	100 with multiple system atrophy, 100 with Parkinson's disease	airborne FA exposure, contact with FA during work-related activities, exposure predates beginning of the study	advanced glial changes, cell loss and depigmentation in substantia nigra, development of multiple system atrophy; possible development of Parkinson's disease
He et al. (2017)[45]	N2a cell line	cells without any treatment, cells with DMSO pre-treatment and FA treatment	Cells with FA treatment, Cells with resveratrol treatment, Cells with FA and resveratrol treatment	incubation in FA	smaller and atrophied cells treated with FA, cell borders were irregular and cellular processes were reduced, FA induced neuronal cell death and tau hyperphosphorylation
Lee et al. (2020)[46]	human	40	50 with ALS	endogenous FA	significantly higher levels of FA detected in plasma of people affected with ALS
Liu et al. (2018)[47]	mouse	10	30 10 – FA exposure: 0.155 mg/kg/day, 10 – FA exposure: 1.55 mg/kg/day, 10 – FA exposure: 15.5 mg/kg/day	formalin intranasal instillation, daily for 7 days	impaired spatial memory in mice exposed to FA concentration 1.55 and 15.5 mg/kg/day, blood–brain barrier (BBB) damage, swollen and deformed hippocampal cells, with loose and disordered cell arrangement, disappearing apical dendrites, decrease in Nissl substance. damaged circular structure of the synaptic glomerulus in the olfactory bulb, FA-induced decrease in the number of cells in cerebral and prefrontal cortex, markedly increased expression of Tau-P in cerebral cortex, increased activation of glia and astrocytes, increased OS level and inflammation
Lu et al. (2013)[48]	mouse, N2a cell, rat hippocampal neuron culture	untreated cells	not disclosed	mice injected with FA solution, cells cultured with 0.5 mM FA solution; hippocampal neuron cultures treated with 0.2 mM FA at the indicated intervals	shrunken cells and atrophied largely processes, hyperphosphorylation and polymerization of tau protein in cytoplasm and nuclei of observed cells and in mice brains, possible role in AD development and other tauopathies
Nie et al. (2007)[49]	SH-SY5Y cell line, rat hippocampal cells	not applicable not applicable	not applicable not applicable	incubation with increasing concentrations of FA from 0.01% to 0.5%	induction of aggregation of human tau protein, possible role in development of tauopathies
Tong et al. (2011)[50]	human	38 patients, autopsy hippocampus tissues from 4 healthy patients, mice	141 patients with dementia, 49 hypertension or diabetes patients, autopsy hippocampus tissues from 4 patients with AD, age-related AD-like animal models – mice	not disclosed	in dementia patients and patients with mild cognitive impairment: higher levels of FA in urine, in the hippocampal tissues from patients with AD and in brains from AD-like mice models: higher levels of FA detected, in mice treated with FA for 30 days: impaired spatial reference memory function
Tong et al. (2015)[51]	human	not disclosed	236 – clinical participants 139 – healthy volunteers	aging related-endogenous FA	increased FA levels in urine of clinical patients, formaldehyde levels showed positive correlation with age and levels of cognitive decline in individuals

Table 3. Influence of formaldehyde on the development of neurodegenerative diseases – cont.

Reference	Model (human/ animal/cell line)	Size – control group	Size – experimental group(s)	Type and duration of exposure	Symptoms, results of study
Tong et al. (2015)[51]	human hippocampal tissue	8 from healthy patients	9 patients with AD	endogenous FA	increased formaldehyde levels and reduced global methylation in hippocampus of people affected with AD, significant decrease of DNMT1 and DNMT3a in the samples from older advanced age group; 1000% increase of levels of hippocampal formaldehyde in AD patients
Tong et al. (2015)[51]	rat	8	8 – recent spatial memory training, 8 – remote spatial memory training	FA injection, once a day for 7 days	deficits in spatial memory of formaldehyde-injected rats, decline in DMNT expression, significant decrease of DNMT1 and DNMT3a in the samples from 30-month-old rats compared to 3-month-old rats
Tong et al. (2015)[51]	SH-SY5Y cell line	not applicable	not applicable	incubation in FA	reduction in DNMT1 and DNMT3a expression levels
Weisskopf et al. (2009)[52]	human	not disclosed	1,156 patients passed due to ALS	unspecified environmental exposure, including FA	doubled mortality rate for ALS for individuals with documented duration of formaldehyde exposure strong dose-response relationship with number of years of exposure
Yuan et al. (2023)[5]	human	not included in the study	not disclosed	tropospheric FA exposure, indoor FA pollution	higher levels of tropospheric and indoor formaldehyde are linked to a higher occurrence of AD

FA – formaldehyde; ROS – reactive oxygen species; OS – oxidative stress; 8-OHdG – 8-hydroxy-2′-deoxyguanosine; SOD – superoxide dismutase; CAT – catalase; PC12 – pheochromocytoma cell line; C6 – rat glial tumor cell line; MDA – malondialdehyde; BDNF – brain-derived neurotrophic factor; HT22 – mouse hippocampal neuronal cell line; Cyt-C – cytochrome C; NO – nitric oxide; H_2S – hydrogen sulfide; CBS – cystathionine β-synthase; GSH – glutathione; AChE – acetylcholinesterase; NMDA – N-methyl-D-aspartate; SH-SY5Y – human neuroblastoma cell line; ALS – amyotrophic lateral sclerosis; Tau-P – hyperphosphorylated tau protein; DNMT1 – DNA (cytosine-5)-methyltransferase 1; DNMT3a – DNA (cytosine-5)-methyltransferase 3 alpha; DMSO – dimethyl sulfoxide; N2a – mouse neuroblastoma cell line; AD – Alzheimer's disease.

levels of 8-hydroxy-deoxyguanosine (8-OHdG), another important marker of OS,[57] in both brain tissue and urine.[17]

Exposure to FA decreases glutathione levels while increasing concentrations of 4-hydroxynonenal.[21] Concentrations of lipid peroxides and oxidized proteins also increased,[16,18,21] as did total ROS levels.[21,22] Increased ROS levels reduced the viability of PC12 cells.[22] This overall increase in OS marker levels following exposure suggests substantial intensification of OS induced by FA.

The intensification of OS may result from multiple pathways. Incubation of PC12 cells with FA leads to a significant decrease in the concentration and production of the antioxidant molecule H_2S. Exposed cells also displayed a significant increase in inducible nitric oxide synthase (iNOS)[16] and neuronal nitric oxide synthase (nNOS) levels and overproduced nitric oxide (NO), which is a major mechanism inhibiting cystathionine β-synthase, the H_2S-producing enzyme in PC12 cells.[22] Formaldehyde also interacts with glutathione (GSH), thereby depleting the organism's antioxidant defenses.[23] Intensification of OS may additionally result from inhibition of paraoxonase-1, an antioxidant enzyme.[58]

It is well established that neurons are highly susceptible to damage caused by OS, especially those located in the hippocampus and amygdala.[59] Given the information presented above, OS may be considered one of the major mechanisms of cellular damage associated with FA exposure.

Apoptosis induced by formaldehyde

Cells exposed to FA showed changes typical of apoptosis.[16,22] Following exposure, neurons in the frontal cortex and hippocampus of rats exhibited extensive degenerative changes, dark pyknotic nuclei, and shrunken cytoplasm.[18,21] Moreover, their axons retracted, and the cells became separated from one another.[21] The cerebellum is also affected by FA exposure, leading to neuronal damage in all its layers, with Purkinje cells being the most severely affected.[16] Formaldehyde exposure also results in altered expression of proteins involved in apoptosis and cellular protection. In capillaries supplying Purkinje cells, Bax protein levels increased. Bax is a pro-apoptotic protein and a member of the *Bcl-2* gene family involved in p53-mediated apoptosis activation.[16]

In PC12 cells, FA exposure induced cytotoxic and apoptotic effects, including decreased levels of Bcl-2 protein and increased release of cytochrome C.[58] Bcl-2 is a protein responsible for inhibiting apoptosis,[60] whereas increased release of cytochrome C into the cytoplasm leads to apoptosome formation and activation of caspase-9, thereby initiating the caspase cascade and inducing apoptosis.[61]

Formaldehyde may also induce ferroptosis – an iron-dependent form of programmed cell death – in HT22 cells (a mouse hippocampal cell line) by increasing expression of genes involved in the ferroptosis pathway. A possible

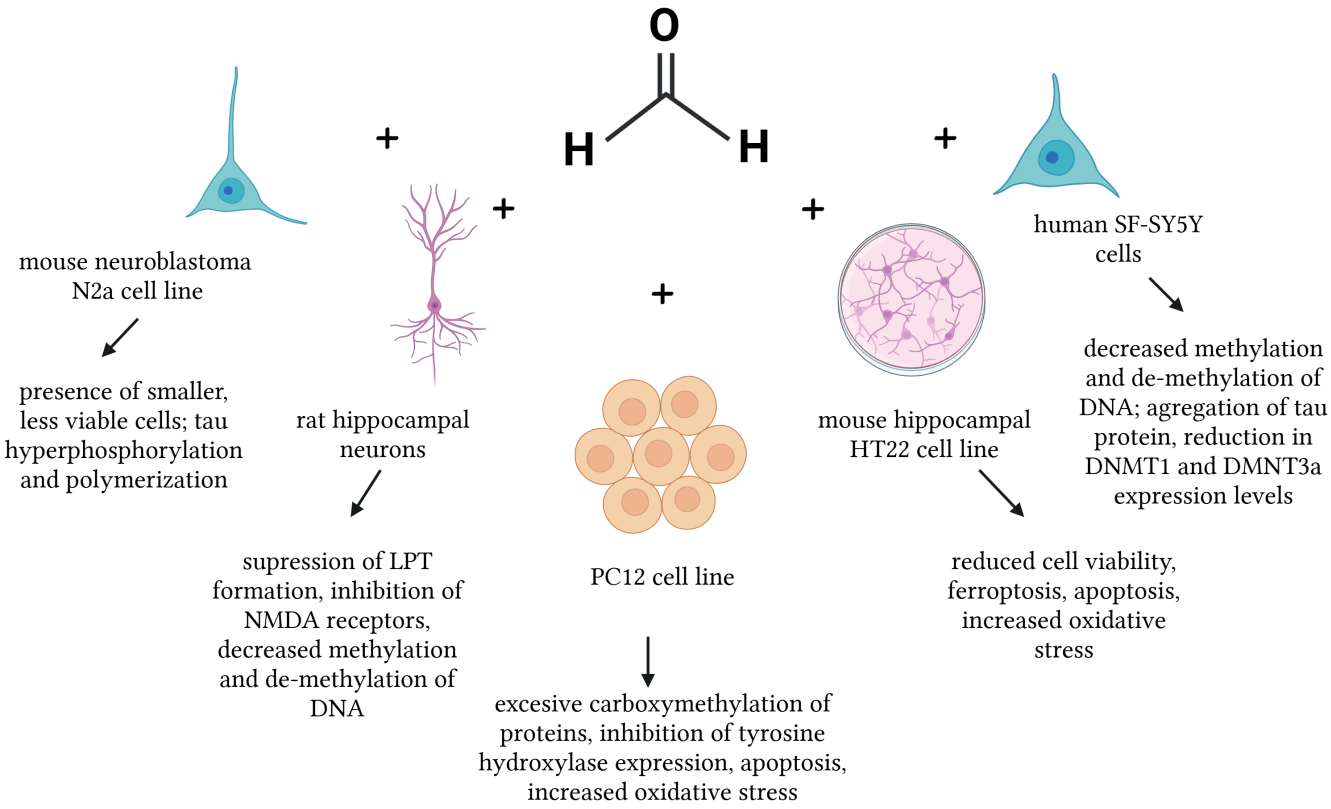


Fig. 2. The range of responses of human and rat/mouse neuronal cell lines to formaldehyde (FA) exposure (created using BioRender.com)

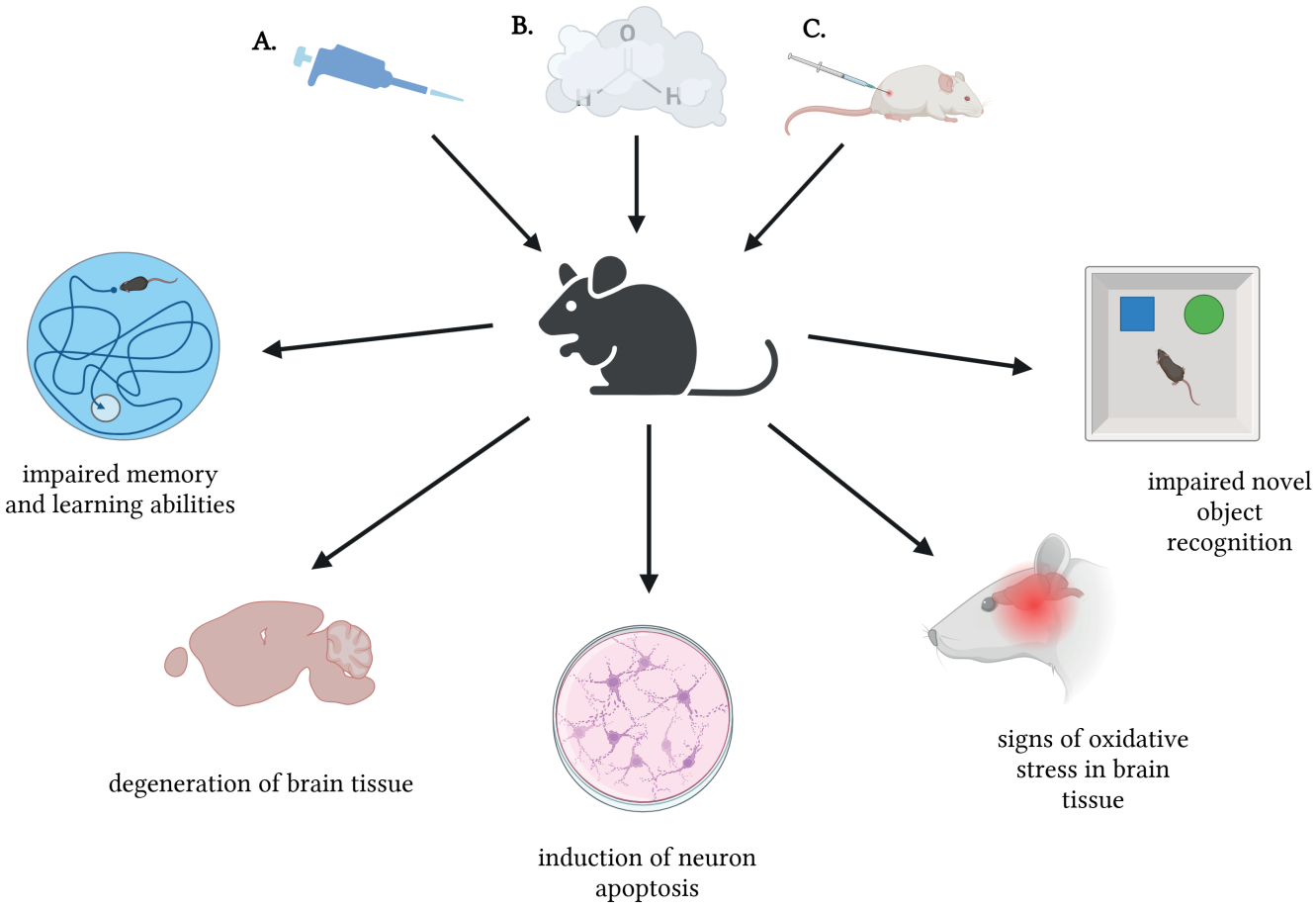


Fig. 3. The effects of formaldehyde (FA) on the function and condition of nervous tissue in mice and rats. Letter labels indicate the studied routes of FA exposure in animals. A. Intranasal instillation of FA; B. Inhalation of FA vapors; C. Intravenous administration of FA solution (created using BioRender.com)

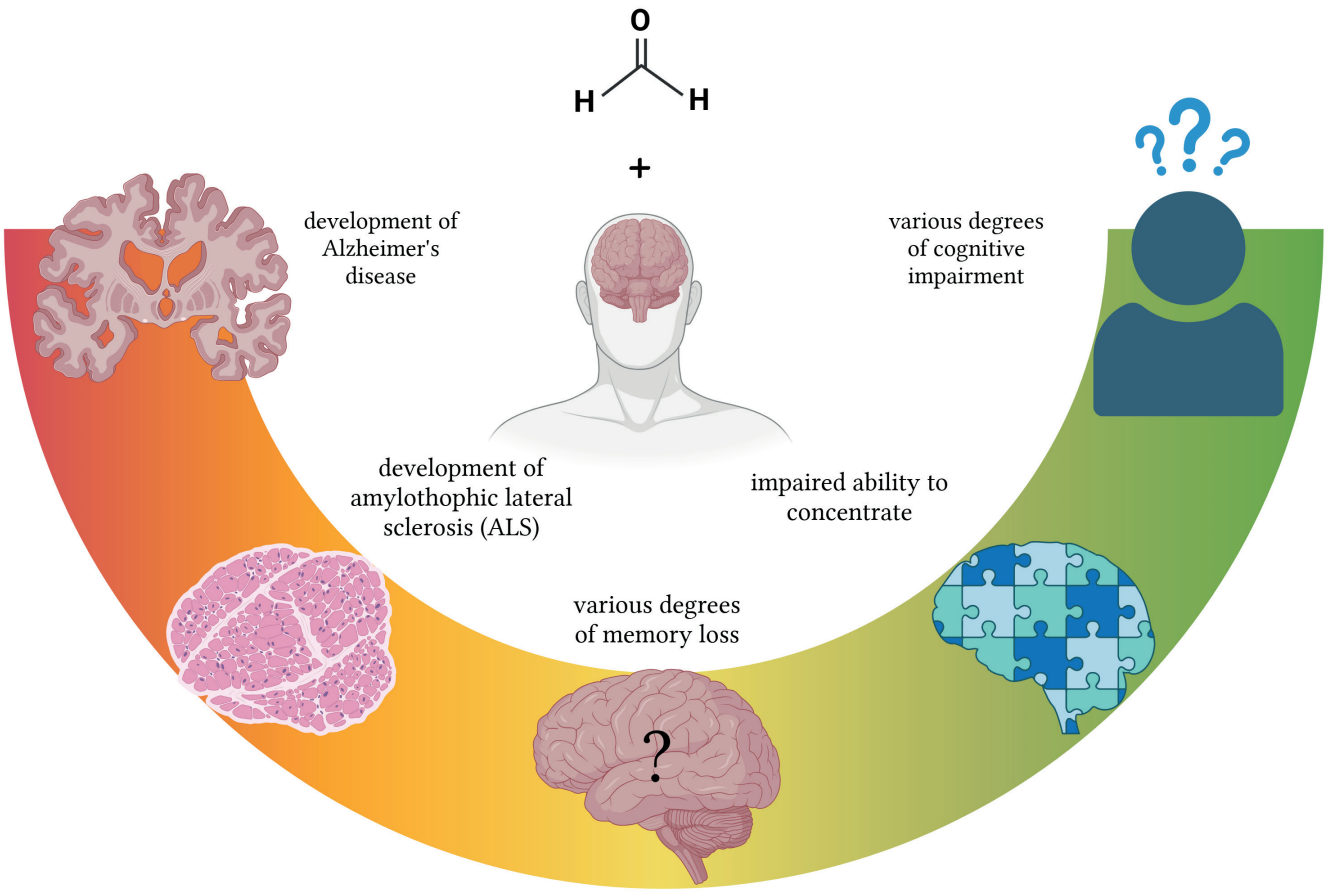


Fig. 4. Potential effects of formaldehyde (FA) on human brain function (created using BioRender.com)

mechanism underlying ferroptosis in this cell line may involve upregulation of the Warburg effect following FA exposure.[21]

Other possible mechanisms of formaldehyde neurotoxicity

Formaldehyde can cross-link proteins with nucleic acids[62] and form covalent bonds with proteins.[19] The presence of covalent bonds between FA and proteins results in accumulation of protein methylation changes typically observed in aging cells.[19] Additionally, cross-linking between proteins and nucleic acids may lead to mutations in the cellular genome.[62] Reduced dopamine activity has also been observed following FA exposure, possibly as a result of decreased expression of tyrosine hydroxylase, an enzyme involved in dopamine synthesis.[38]

Workers exposed to FA show a significant correlation between occupational FA exposure and increased blood acetylcholinesterase (AChE) activity,[24] which is used as a marker of neurotoxicity.[63] Therefore, elevated AChE activity may indicate neuronal damage induced by FA.

Neuronal death in the hippocampus disrupts learning processes and memory.[64] Damage to cerebellar neurons is associated with impaired motor coordination, imbalance, and loss of spatial orientation.[65] Finally, damage to the frontal lobes may lead to personality changes, reduced attention span, impaired concentration, language dysfunction, and impaired decision-making.[66] These examples demonstrate how FA-induced neuronal damage may contribute to cognitive decline.

Formaldehyde exposure leads to cognitive impairment and behavioral changes

The correlation between FA exposure and the occurrence of numerous types of cognitive impairment in humans suggests that FA may be harmful to human health (Fig. 4).[7,14] Even moderate but persistent exposure to FA may significantly affect the cognitive functions of exposed individuals.[31] People occupationally exposed to FA had a statistically higher risk of cognitive decline. The risk of impairment increased with the duration of exposure. Moreover, FA-induced changes persisted even after exposure had ended.[31]

Inhalation of airborne FA in the form of "fry joints" or cigarettes immersed in FA solution and laced with phencyclidine (PCP) was associated with decreased language abilities and vocabulary decline.[36] In the case of fry joints, however, it remains unclear whether the observed decline resulted solely from FA exposure or whether other components of "fry" also contributed.[36] Studies have also

reported impaired abstract thinking following FA exposure.[36] Additional studies demonstrated similar language impairments after occupational exposure.[31] Another important consequence of FA exposure is reduced attention span and, in some cases, impaired concentration.[31,40] Some individuals exposed to FA also exhibited reduced affect and psychomotor slowing.[31,40]

Chemical analyses of workers in melamine tableware factories showed that individuals exposed to airborne FA had significantly elevated blood levels of active AChE. Such alterations are thought to contribute to cognitive impairment through reduction of cholinergic signaling in neurons.[24] Cases of acute airborne FA exposure leading to neurobehavioral impairment and cognitive decline have also been described, including confusion, impaired concentration, and deficits in sequential processing. Depression and emotional distress accompanied these symptoms, persisting for up to 3 years after exposure.[38]

Studies in rat models also demonstrated that animals exposed to FA exhibited depression-like behaviors, such as reduced locomotor activity, decreased food and water intake, and anxiety-like behavior.[33] Rats exposed to FA for short periods (3 h/day for 2 days in 3 groups receiving 5 ppm, 10 ppm, and 20 ppm FA[67]; or 2 h/day for 10 days in 3 groups receiving 0.1 ppm, 0.5 ppm, and 5.4 ppm FA35) exhibited reduced motor activity.[35] Prolonged exposure led to enhanced aggressive behavior. This change was associated with disruption of dopaminergic and serotonergic pathways in the frontal cortex. Reduced activity of dopamine and serotonin neurons results in impaired concentration and attention deficits. Moreover, FA has been shown to reduce melatonin levels in rats. Rodents examined in the study conducted by Mei et al. demonstrated memory impairment and abnormal behaviors, including remaining motionless in a huddled position and decreased active exploration.[37] Lower melatonin levels may be associated with sleep disorders reported after FA exposure.[54] Studies have shown that decreased melatonin levels are also associated with depression, anxiety, Alzheimer's disease (AD), and other types of dementia. Melatonin supplementation has been shown to attenuate the effects of FA-induced OS and structural hippocampal damage, thereby reducing cognitive decline.[37]

Formaldehyde exposure leads to memory loss

Formaldehyde can have a detrimental impact on memory. Several studies have investigated this issue, mostly using animal models, although some also involved humans. Animal studies investigating memory were conducted in rats and mice. These experiments employed various tests specifically designed to assess cognitive function in animals following exposure to specific factors, in this case FA in gaseous form or in solutions of various concentrations. Animals exposed to FA, either in gaseous form or through intracerebroventricular injection, exhibited memory alterations characterized by impaired learning ability, as demonstrated in the Morris water maze and water labyrinth tests,[34,35,41] as well as decreased memory performance observed in the step-down test.[26] Studies have shown that FA exposure can reduce spatial memory performance, as demonstrated in the Morris water maze and dry food-reward maze tests.[20,32,39,42] Impaired novel object recognition may also indicate memory dysfunction and has been observed in mice exposed to FA.[20,32]

Microscopic and biochemical studies have also been performed to investigate the possible mechanisms through which FA induces memory impairment. Formaldehyde may disrupt memory-related processes in the hippocampus by inhibiting alcohol dehydrogenase-3 and aldehyde dehydrogenase-2, thereby suppressing long-term potentiation (LTP) formation, or by nonspecifically inhibiting N-methyl-D-aspartate (NMDA) receptors,[43] both of which are believed to be associated with memory formation.[68] Formaldehyde may also directly impair the brain's ability to acquire new information by inducing hippocampal neuronal death,[18,37] as demonstrated in rats following intraperitoneal injections of FA at a dose of 10 mg/kg for 10 days.[18]

Studies investigating the influence of FA on memory in humans included a total of 944 participants. Most focused on occupational exposure.[25,27–30] One study was based on full-time residents of houses insulated with urea-formaldehyde foam.[40] Only 1 study was conducted under experimental conditions, in which participants were placed in a climate chamber and exposed to gaseous FA at concentrations that increased throughout the experiment.[3] Individuals exposed to FA appeared to exhibit deficits in short-term[25] and long-term memory,[28,29] as well as visual memory, history memory,[27] and episodic verbal memory.[31] Formaldehyde originating from building materials such as foam insulation may also impair memory processes.[40] Finally, FA exposure may cause memory loss in affected individuals.[30] Acute and severe exposure to FA may impair both general memory and the ability to acquire new information.[38]

Neurodegenerative diseases occurrence after formaldehyde exposure

Exposure to FA is not limited to healthcare or laboratory workers; individuals living in large urban areas may also be at risk. In major metropolitan areas, air pollution, including FA, may adversely affect health.[4] Elevated levels of antibodies targeting FA-bound protein adducts have been detected in children living in urban environments exposed to air pollution. We would like to highlight a possible link between this heightened immune response and an increased risk of neurodegenerative diseases such as AD or Parkinson's disease.[4] Occupational exposure to FA has also been suggested to influence the development of Parkinson's disease.[38]

Alzheimer's disease in relation to formaldehyde

High tropospheric and indoor FA levels may be associated with an increased risk of AD in humans.[5] Mice exposed to high concentrations of FA (1.55 mg/kg/day and 15.5 mg/kg/day) through nasal instillation exhibit brain changes resembling early manifestations of AD. These mice show disruption of the BBB, inflammation, and increased production of ROS, leading to OS. Additionally, degenerative structural changes are observed in the hippocampus, olfactory bulb, cerebral cortex, and prefrontal cortex. Exposed mice also display impaired spatial memory acquisition and retention, as demonstrated in the Morris water maze test.[47] These findings have been proposed as potential mechanisms underlying AD.[69]

Formaldehyde levels in urine and the hippocampus increase with age and cognitive decline in both rats and humans with AD. In the hippocampus, elevated FA levels correlate with reduced expression of DNA methyltransferase-3 (DNMT3), resulting in suppression of its activity, as well as decreased expression of DNA methyltransferase-1 (DNMT1), as demonstrated in studies of human hippocampal tissue and rats treated with FA injections. These enzymes are responsible for DNA methylation; consequently, DNA methylation was also reduced under these conditions.[51] DNA methylation is crucial for spatial memory formation and maintenance of long-term memory.[70] Inhibition of DNMT activity and expression by FA may therefore contribute to the pathophysiology of AD.[51]

Elevated FA levels in mouse and rat cells have also been associated with the presence of hyperphosphorylated tau protein analogs linked to AD.[47,48] Formaldehyde particularly affects nuclear tau protein, inducing hyperphosphorylation mediated by glycogen synthase kinase-3β (GSK-3β), as demonstrated in studies of neuroblastoma (N2a) cell nuclei and mouse brains.[48] This hyperphosphorylation is accompanied by aggregation of reversible hyperphosphorylated tau protein polymers within the cytoplasm,[48] as well as morphological changes in these cells, including shrinkage, disruption of cellular processes,[45,48] and ultimately cell death.[45]

Formaldehyde also influences tau protein misfolding and aggregation, which may contribute to the development of tauopathies. Tau protein plays a key role in maintaining cytoskeletal structure through its association with microtubules.[71] Formaldehyde induces formation of amyloid-like aggregates in rats and human neuroblastoma cells,[49] amyloid plaques in mice and rats,[47] and amyloid-β accumulation in mice.[50]

Alzheimer's disease is the most common cause of dementia.[72] Formaldehyde levels in the hippocampi of individuals with dementia are significantly elevated. However, the study sample was small (n = 4); therefore, further studies are needed to confirm or refute this observation. Urine samples demonstrated significantly increased FA levels in patients with AD and slightly increased levels in patients with mild cognitive impairment. Studies using transgenic mice have shown that elevated FA levels accompany amyloid-β deposition.[50] Formaldehyde increases both the number and size of amyloid-β oligomers.[52] The potential use of urinary FA measurement as a biomarker of dementia cannot be excluded.[50]

Amyotrophic lateral sclerosis and formaldehyde

Among individuals exposed to FA, the mortality rate from amyotrophic lateral sclerosis (ALS) is more than twice as high as that observed in non-exposed individuals, and the duration of FA exposure appears to have a significant impact.[73] Elevated plasma FA levels may increase the risk of ALS or exacerbate disease progression. Plasma FA levels in patients with ALS were higher than those observed in healthy individuals, as demonstrated in the study conducted by Lee et al.[46] Increased FA levels induce abnormal folding and aggregation of tau protein in nerve cells, leading to apoptosis, as shown in the study by Nie et al. conducted on SH-SY5Y human neuroblastoma cells.[49] Additionally, the presence of FA in cells increases mitochondrial membrane permeability and decreases SOD activity, thereby causing oxidative damage, as observed by Hanna et al.[44] These are all mechanisms associated with ALS.

Air pollution, including the presence of FA, has also been suggested to influence the development of multiple system atrophy. This was supported by findings of extensive cell loss, glial alterations such as glial cytoplasmic inclusions, and depigmentation in the substantia nigra (SN) and locus coeruleus (LC), identified post mortem in a previously exposed patient.[38]

Limitations of the study

However precise and thorough we aimed to be, our review is not without limitations. First, the heterogeneity of the studies available on FA exposure-related diseases made any type of meta-analysis impossible. We also acknowledge that our search strategy, although broad, may have lacked certain details, which could have resulted in omission of some relevant studies. Even though the data included in this review were collected using structured tables with predefined variables of interest, some information may still have been overlooked; e.g., certain important details may not have been fully captured.

Conclusions

The purpose of this study was to gather and analyze the available literature describing the influence of FA on the human brain and cognition. We aimed to highlight

potential problems that may result from exposure to this compound. Formaldehyde may negatively affect cognitive function. Reported symptoms include language and vocabulary difficulties, reduced attention span, depression, emotional distress, and anxiety. Sleep disturbances have also been described following FA exposure. Most studies in this area focus on memory and demonstrate substantial impairment of memory function. Formaldehyde also promotes pathological changes associated with neurodegenerative diseases such as AD and ALS. At the cellular level, FA primarily disrupts the function of the frontal cortex, hippocampus, and cerebellum. The main mechanisms underlying neuronal damage and death appear to be OS and apoptosis induction. Based on the available evidence, FA is not neutral to CNS health. However, most studies have been conducted using animal models. Therefore, further research involving human participants is needed and may be feasible due to the widespread use of FA in industry and medicine.

This review provides additional arguments supporting the need to identify safer alternatives to FA for tissue preservation, particularly in medicine. We hope that increasing awareness of the effects of FA on brain function will contribute to the development of improved strategies for minimizing exposure or, where possible, eliminating it entirely. This issue is especially important in anatomy, where FA remains the primary preservative used for cadavers. As a result, students and teachers are routinely exposed to chemical agents that may impair concentration and learning ability, in addition to potentially increasing the risk of serious neurological disorders. It is in our best interest to create an environment in which young people can study safely and develop into highly qualified specialists and physicians.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Mateusz Drążyk https://orcid.org/0009-0004-6594-7112
Zuzanna Pyc https://orcid.org/0009-0007-9477-2122
Szymon J. Pietrzyk https://orcid.org/0009-0008-4344-6586
Antonina Gajda-Janiak https://orcid.org/0009-0008-7442-0768
Filip Godziszewski https://orcid.org/0009-0007-3198-4822
Oliwier Pioterek https://orcid.org/0009-0000-7896-9266
Michał Tulski https://orcid.org/0009-0001-3867-5850
Mateusz Mazurek https://orcid.org/0009-0008-0066-6087
Zygmunt A. Domagała https://orcid.org/0000-0002-2317-1932

References

1. Musiał A, Gryglewski RW, Kielczewski S, Loukas M, Wajda J. Formalin use in anatomical and histological science in the 19th and 20th centuries. *Folia Med Cracov*. 2016;56(3):31–40. PMID:28275269.
2. Brenner E. Human body preservation: Old and new techniques. *J Anat*. 2014;224(3):316–344. doi:10.1111/joa.12160
3. Yahyaei E, Majlesi B, Naimi Joubani M, et al. Occupational exposure and risk assessment of formaldehyde in the pathology departments of hospitals. *Asian Pac J Cancer Prev*. 2020;21(5):1303–1309. doi:10.31557/APJCP.2020.21.5.1303
4. Calderón-Garcidueñas L, Vojdani A, Blaurock-Busch E, et al. Air pollution and children: Neural and tight junction antibodies and combustion metals, the role of barrier breakdown and brain immunity in neurodegeneration. *J Alzheimers Dis*. 2014;43(3):1039–1058. doi:10.3233/JAD-141365
5. Yuan Y, Wu Y, Zhao H, et al. Tropospheric formaldehyde levels infer ambient formaldehyde-induced brain diseases and global burden in China, 2013–2019. *Sci Total Environ*. 2023;883:163553. doi:10.1016/j.scitotenv.2023.163553
6. Nelson PF, Tibbett AR, Day SJ. Effects of vehicle type and fuel quality on real world toxic emissions from diesel vehicles. *Atmospheric Environment*. 2008;42(21):5291–5303. doi:10.1016/j.atmosenv.2008.02.049
7. Suarez-Bertoa R, Selleri T, Gioria R, et al. Real-time measurements of formaldehyde emissions from modern vehicles. *Energies*. 2022;15(20):7680. doi:10.3390/en15207680
8. Baker RR. The generation of formaldehyde in cigarettes: Overview and recent experiments. *Food Chem Toxicol*. 2006;44(11):1799–1822. doi:10.1016/j.fct.2006.05.017
9. Soltanpour Z, Mohammadian Y, Fakhri Y. The exposure to formaldehyde in industries and health care centers: A systematic review and probabilistic health risk assessment. *Environ Res*. 2022;204:112094. doi:10.1016/j.envres.2021.112094
10. Doi T, Kajimura K, Taguchi S. Survey of formaldehyde (FA) concentration in cosmetics containing FA-donor preservatives. *J Health Sci*. 2010;56(1):116–122. doi:10.1248/jhs.56.116
11. Malinauskiene L, Blaziene A, Chomiciene A, Isaksson M. Formaldehyde may be found in cosmetic Productseven when unlabelled. *Open Med (Wars)*. 2015;10(1):323–328. doi:10.1515/med-2015-0047
12. Pontel LB, Rosado IV, Burgos-Barragan G, et al. Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen. *Mol Cell*. 2015;60(1):177–188. doi:10.1016/j.molcel.2015.08.020
13. Cogliano V, Grosse Y, Baan R, Straif K, Secretan B, El Ghissassi F; WHO International Agency for Research on Cancer. Advice on formaldehyde and glycol ethers. *Lancet Oncol*. 2004;5(9):528. doi:10.1016/s1470-2045(04)01562-1
14. Protano C, Buomprisco G, Cammalleri V, et al. The carcinogenic effects of formaldehyde occupational exposure: A systematic review. *Cancers (Basel)*. 2021;14(1):165. doi:10.3390/cancers14010165
15. Rana I, Rieswijk L, Steinmaus C, Zhang L. Formaldehyde and brain disorders: A meta-analysis and bioinformatics approach. *Neurotox Res*. 2021;39(3):924–948. doi:10.1007/s12640-020-00320-y
16. Askar EM, Halloull NM. Formaldehyde-induced neurotoxicity in rat cerebellar cortex and possible protective effects of fatty acids from omega 3 and wheat germ oil supplement: A histopathological and biochemical study. *J Histotechnol*. 2018;41(3):79–87. doi:10.1080/01478885.2018.1458176
17. Ciftci G, Aksoy A, Cenesiz S, et al. Therapeutic role of curcumin in oxidative DNA damage caused by formaldehyde. *Microsc Res Tech*. 2015;78(5):391–395. doi:10.1002/jemt.22485
18. Gurel A, Coskun O, Armutcu F, Kanter M, Ozen OA. Vitamin E against oxidative damage caused by formaldehyde in frontal cortex and hippocampus: Biochemical and histological studies. *J Chem Neuroanat*. 2005;29(3):173–178. doi:10.1016/j.jchemneu.2005.01.001
19. Lee ES, Chen H, Hardman C, Simm A, Charlton C. Excessive S-adenosyl-l-methionine-dependent methylation increases levels of methanol, formaldehyde and formic acid in rat brain striatal homogenates: Possible role in S-adenosyl-l-methionine-induced Parkinson's disease-like disorders. *Life Sci*. 2008;83(25–26):821–827. doi:10.1016/j.lfs.2008.09.020
20. Li X, Zhuang YY, Wu L, et al. Hydrogen sulfide ameliorates cognitive dysfunction in formaldehyde-exposed rats: Involvement in the upregulation of brain-derived neurotrophic factor. *Neuropsychobiology*. 2020;79(2):119–130. doi:10.1159/000501294
21. Li XN, Yang SQ, Li M, et al. Formaldehyde induces ferroptosis in hippocampal neuronal cells by upregulation of the Warburg effect. *Toxicology*. 2021;448:152650. doi:10.1016/j.tox.2020.152650
22. Tang XQ, Fang HR, Zhou CF, et al. A novel mechanism of formaldehyde neurotoxicity: Inhibition of hydrogen sulfide generation by promoting overproduction of nitric oxide. *PLoS One*. 2013;8(1):e54829. doi:10.1371/journal.pone.0054829

23. Umansky C, Morellato AE, Rieckher M, et al. Endogenous formaldehyde scavenges cellular glutathione resulting in redox disruption and cytotoxicity. *Nat Commun*. 2022;13(1):745. doi:10.1038/s41467-022-28242-7
24. Zendehdel R, Fazli Z, Mazinani M. Neurotoxicity effect of formaldehyde on occupational exposure and influence of individual susceptibility to some metabolism parameters. *Environ Monit Assess*. 2016;188(11):648. doi:10.1007/s10661-016-5662-z
25. Bach B, Pedersen OF, Molhave L. Human performance during experimental formaldehyde exposure. *Environ Int*. 1990;16:105–113. doi:10.1016/0160-4120(90)90150-5
26. Dou XJ, Zhang Y, Wu YH. Study on neurotoxicity of formaldehyde in mice. *Toxicol Environ Health Sci*. 2012;4(2):115–120. doi:10.1007/s13530-012-0125-6
27. Kilburn KH, Warshaw R, Thornton JC. Formaldehyde impairs memory, equilibrium, and dexterity in histology technicians: Effects which persist for days after exposure. *Arch Environ Health*. 1987;42(2):117–120. doi:10.1080/00039896.1987.9935806
28. Kilburn KH, Warshaw R, Boylen CT, et al. Toxic effects of formaldehyde and solvents in histology technicians: A preliminary report. *J Histotechnol*. 1983;6(2):73–76. doi:10.1179/his.1983.6.2.73
29. Kilburn KH, Warshaw R, Boylen CT, et al. Pulmonary and neurobehavioral effects of formaldehyde exposure. *Arch Environ Health*. 1985;40(5):254–260. doi:10.1080/00039896.1985.10545928
30. Kilburn KH, Seidman BC, Warshaw R. Neurobehavioral and respiratory symptoms of formaldehyde and xylene exposure in histology technicians. *Arch Environ Health*. 1985;40(4):229–233. doi:10.1080/00039896.1985.10545924
31. Letellier N, Gutierrez LA, Pilorget C, et al. Association between occupational exposure to formaldehyde and cognitive impairment. *Neurology*. 2022;98(6):e633–e640. doi:10.1212/WNL.0000000000013146
32. Li F, Yujie Qin, Gong S, Zhang H, Ding S. Learning and memory impairment of mice caused by gaseous formaldehyde. *Environ Res*. 2020;184:109318. doi:10.1016/j.envres.2020.109318
33. Li Y, Song Z, Ding Y, et al. Effects of formaldehyde exposure on anxiety-like and depression-like behavior, cognition, central levels of glucocorticoid receptor and tyrosine hydroxylase in mice. *Chemosphere*. 2016;144:2004–2012. doi:10.1016/j.chemosphere.2015.10.102
34. Lu Z, Li CM, Qiao Y, Yan Y, Yang X. Effect of inhaled formaldehyde on learning and memory of mice. *Indoor Air*. 2008;18(2):77–83. doi:10.1111/j.1600-0668.2008.00524.x
35. Malek FA, Möritz KU, Fanghänel J. A study on the effect of inhalative formaldehyde exposure on water labyrinth test performance in rats. *Ann Anat*. 2003;185(3):277–285. doi:10.1016/S0940-9602(03)80040-7
36. Marceaux JC, Dilks LS, Hixson S. Neuropsychological effects of formaldehyde use. *J Psychoactive Drugs*. 2008;40(2):207–210. doi:10.1080/02791072.2008.10400632
37. Mei Y, Duan C, Li X, et al. Reduction of endogenous melatonin accelerates cognitive decline in mice in a simulated occupational formaldehyde exposure environment. *Int J Environ Res Public Health*. 2016;13(3):258. doi:10.3390/ijerph13030258
38. Perna RB, Bordini EJ, Deinzer-Lifrak M. A case of claimed persistent neuropsychological sequelae of chronic formaldehyde exposure: Clinical, psychometric, and functional findings. *Arch Clin Neuropsychol*. 2001;16(1):33–44. PMID:14590191.
39. Pitten FA, Kramer A, Herrmann K, Bremer J, Koch S. Formaldehyde neurotoxicity in animal experiments. *Pathol Res Pract*. 2000;196(3):193–198. doi:10.1016/S0344-0338(00)80100-4
40. Schenker MB, Weiss ST, Murawski BJ. Health effects of residence in homes with urea formaldehyde foam insulation: A pilot study. *Environ Int*. 1982;8(1–6):359–363. doi:10.1016/0160-4120(82)90050-2
41. Tang XQ, Zhuang YY, Zhang P, et al. Formaldehyde impairs learning and memory involving the disturbance of hydrogen sulfide generation in the hippocampus of rats. *J Mol Neurosci*. 2013;49(1):140–149. doi:10.1007/s12031-012-9912-4
42. Tong Z, Han C, Luo W, et al. Aging-associated excess formaldehyde leads to spatial memory deficits. *Sci Rep*. 2013;3(1):1807. doi:10.1038/srep01807
43. Tong Z, Han C, Luo W, et al. Accumulated hippocampal formaldehyde induces age-dependent memory decline. *Age (Dordr)*. 2013;35(3):583–596. doi:10.1007/s11357-012-9388-8
44. Hanna PA, Jankovic J, Kirkpatrick JB. Multiple system atrophy: The putative causative role of environmental toxins. *Arch Neurol*. 1999;56(1):90. doi:10.1001/archneur.56.1.90
45. He X, Li Z, Rizak JD, et al. Resveratrol attenuates formaldehyde induced hyperphosphorylation of Tau protein and cytotoxicity in N2a cells. *Front Neurosci*. 2017;10:598. doi:10.3389/fnins.2016.00598
46. Lee A, Arachchige BJ, Reed S, Henderson R, Aylward J, McCombe PA. Plasma from some patients with amyotrophic lateral sclerosis exhibits elevated formaldehyde levels. *J Neurol Sci*. 2020;409:116589. doi:10.1016/j.jns.2019.116589
47. Liu X, Zhang Y, Wu R, et al. Acute formaldehyde exposure induced early Alzheimer-like changes in mouse brain. *Toxicol Mech Methods*. 2018;28(2):95–104. doi:10.1080/15376516.2017.1368053
48. Lu J, Miao J, Su T, Liu Y, He R. Formaldehyde induces hyperphosphorylation and polymerization of Tau protein both in vitro and in vivo. *Biochim Biophys Acta Gen Subj*. 2013;1830(8):4102–4116. doi:10.1016/j.bbagen.2013.04.028
49. Nie CL, Wang XS, Liu Y, Perrett S, He RQ. Amyloid-like aggregates of neuronal tau induced by formaldehyde promote apoptosis of neuronal cells. *BMC Neurosci*. 2007;8(1):9. doi:10.1186/1471-2202-8-9
50. Tong Z, Zhang J, Luo W, et al. Urine formaldehyde level is inversely correlated to mini mental state examination scores in senile dementia. *Neurobiol Aging*. 2011;32(1):31–41. doi:10.1016/j.neurobiolaging.2009.07.013
51. Tong Z, Han C, Qiang M, et al. Age-related formaldehyde interferes with DNA methyltransferase function, causing memory loss in Alzheimer's disease. *Neurobiol Aging*. 2015;36(1):100–110. doi:10.1016/j.neurobiolaging.2014.07.018
52. Weisskopf MG, Morozova N, O'Reilly EJ, et al. Prospective study of chemical exposures and amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry*. 2009;80(5):558–561. doi:10.1136/jnnp.2008.156976
53. Shcherbakova LN, Tel'pukhov VI, Trenin SO, Bashilov IA, Lapkina TI. Permeability of the blood-brain barrier to intra-arterial formaldehyde [in Russian]. *Biull Eksp Biol Med*. 1986;102(11):573–575. PMID:3779084.
54. Janero DR. Malondialdehyde and thiobarbituric acid-reactivity as diagnostic indices of lipid peroxidation and peroxidative tissue injury. *Free Radic Biol Med*. 1990;9(6):515–540. doi:10.1016/0891-5849(90)90131-2
55. Nandi A, Yan LJ, Jana CK, Das N. Role of catalase in oxidative stress- and age-associated degenerative diseases. *Oxid Med Cell Longev*. 2019;2019:9613090. doi:10.1155/2019/9613090
56. Islam MN, Rauf A, Fahad FI, et al. Superoxide dismutase: An updated review on its health benefits and industrial applications. *Crit Rev Food Sci Nutr*. 2022;62(26):7282–7300. doi:10.1080/10408398.2021.1913400
57. Valavanidis A, Vlachogianni T, Fiotakis C. 8-hydroxy-2'-deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis. *J Environ Sci Health C*. 2009;27(2):120–139. doi:10.1080/10590500902885684
58. Tang X, Ren Y, Chen R, et al. Formaldehyde induces neurotoxicity to PC12 cells involving inhibition of paraoxonase-1 expression and activity. *Clin Exp Pharmacol Physiol*. 2011;38(4):208–214. doi:10.1111/j.1440-1681.2011.05485.x
59. Salim S. Oxidative stress and the central nervous system. *J Pharmacol Exp Ther*. 2017;360(1):201–205. doi:10.1124/jpet.116.237503
60. Hockenbery D, Nuñez G, Milliman C, Schreiber RD, Korsmeyer SJ. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature*. 1990;348(6299):334–336. doi:10.1038/348334a0
61. Ow YLP, Green DR, Hao Z, Mak TW. Cytochrome c: Functions beyond respiration. *Nat Rev Mol Cell Biol*. 2008;9(7):532–542. doi:10.1038/nrm2434
62. Craft TR, Bermudez E, Skopek TR. Formaldehyde mutagenesis and formation of DNA-protein crosslinks in human lymphoblasts in vitro. *Mutat Res*. 1987;176(1):147–155. doi:10.1016/0027-5107(87)90262-4
63. Kalafatakis K, Gkanti V, Mackenzie-Gray Scott C, Zarros A, Baillie G, Tsakiris S. Acetylcholinesterase activity as a neurotoxicity marker within the context of experimentally-simulated hyperprolinaemia: An in vitro approach. *J Nat Sc Biol Med*. 2015;6(3):98. doi:10.4103/0976-9668.166099
64. Anand K, Dhikav V. Hippocampus in health and disease: An overview. *Ann Indian Acad Neurol*. 2012;15(4):239. doi:10.4103/0972-2327.104323

65. Therrien AS, Bastian AJ. Cerebellar damage impairs internal predictions for sensory and motor function. *Curr Opin Neurobiol*. 2015; 33:127–133. doi:10.1016/j.conb.2015.03.013
66. Yu LQ, Kan IP, Kable JW. Beyond a rod through the skull: A systematic review of lesion studies of the human ventromedial frontal lobe. *Cogn Neuropsychol*. 2020;37(1–2):97–141. doi:10.1080/02643294.2019.1690981
67. Boja JW, Nielsen JA, Foldvary E, Truitt EB. Acute low-level formaldehyde behavioural and neurochemical toxicity in the rat. *Prog Neuropsychopharmacol Biol Psychiatry*. 1985;9(5–6):671–674. doi:10.1016/0278-5846(85)90038-7
68. Lynch MA. Long-term potentiation and memory. *Physiol Rev*. 2004; 84(1):87–136. doi:10.1152/physrev.00014.2003
69. Khan S, Barve KH, Kumar MS. Recent advancements in pathogenesis, diagnostics and treatment of Alzheimer's disease. *Curr Neuropharmacol*. 2020;18(11):1106–1125. doi:10.2174/1570159X18666200528142429
70. Feng J, Zhou Y, Campbell SL, et al. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci*. 2010;13(4):423–430. doi:10.1038/nn.2514
71. Schweers O, Schönbrunn-Hanebeck E, Marx A, Mandelkow E. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J Biol Chem*. 1994;269(39): 24290–24297. doi:10.1016/S0021-9258(19)51080-8
72. Chen K, Maley J, Yu PH. Potential implications of endogenous aldehydes in β-amyloid misfolding, oligomerization and fibrillogenesis. *J Neurochem*. 2006;99(5):1413–1424. doi:10.1111/j.1471-4159.2006.04181.x
73. Roberts AL, Johnson NJ, Cudkowicz ME, Eum KD, Weisskopf MG. Job-related formaldehyde exposure and ALS mortality in the USA. *J Neurol Neurosurg Psy*chiatry. 2016;87(7):786–788. doi:10.1136/jnnp-2015-310750

Epidemiology of tuberculosis, scabies, and enteric infections in Polish prisons (2002–2023): A nationwide data analysis and systematic review

Rafał Korkosz[1,A–D], Agata Trzcionka[1,E], Robert Deręgowski[2,C], Maksymilian Kiełbratowski[1,B], Anna Kuśka-Kiełbratowska[3,B], Mansur Rahnama-Hezavah[4,E], Marta Tanasiewicz[1,E,F]

[1] Department of Conservative Dentistry with Endodontics, Faculty of Medical Sciences in Zabrze, Medical University of Silesia, Bytom, Poland
[2] ZF Automotive Systems Poland Sp. z o.o. Czestochowa Electronics Plant, Poland
[3] Department of Periodontal Diseases and Oral Mucosa Diseases, Faculty of Medical Sciences in Zabrze, Poland
[4] Department of Dental Surgery, Medical University of Lublin, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(6):1109–1121

Address for correspondence
Agata Trzcionka
E-mail: atrzcionka@gmail.com

Funding sources
This research received internal academic funding from the Medical University of Silesia (grant No. BNW-1-102/N/5/K).

Conflict of interest
None declared

Received on June 22, 2025
Reviewed on July 26, 2025
Accepted on September 10, 2025

Published online on June 23, 2026

Abstract

The specific conditions prevailing in prisons increase the risk of disease transmission among inmates. Several factors influence the risk of infectious disease transmission in prisons, including overcrowding, limited access to water, delayed diagnosis, and poor ventilation. The aim of this study was to assess the burden of selected infectious diseases among Polish prisoners between 2002 and 2023 and to analyze the literature addressing these diseases published between 2015 and 2025. In the 1st part, a systematic review was conducted. In the 2nd part, the results of the authors' own research were presented. The source material was obtained from Statistics Poland. An increase in tuberculosis (TB) cases was observed from the early to mid-2010s, peaking in 2012. Subsequently, a decrease in TB cases was noted after 2012, reaching a nadir in 2021, followed by a resurgence in 2023. A decrease in scabies cases was also observed. However, from 2009 onward, this trend reversed, with cases increasing and peaking in 2014. Subsequently, the number of cases reached a new low in 2022, before increasing again in 2023. For *Salmonella*/*Shigella*, the number of tests remained around 5,000 in the early years, reaching a peak of 8,876 in 2020. This was followed by a decline, with 5,204 tests recorded in 2023. To minimize the risk of infectious disease transmission in prisons, several preventive measures should be implemented, including screening of newly admitted prisoners, introduction of prophylactic programs, and development of standardized procedures to follow in cases of infection.

Key words: prison health, tuberculosis pulmonary epidemiology, scabies epidemiology, enteric diseases epidemiology, overcrowding

Cite as
Korkosz R, Trzcionka A, Deręgowski R, et al. Epidemiology of tuberculosis, scabies, and enteric infections in Polish prisons (2002–2023): A nationwide data analysis and systematic review. *Adv Clin Exp Med*. 2026;35(6):1109–1121. doi:10.17219/acem/210555

DOI
10.17219/acem/210555

Highlights

- The use of a Python script during the systematic review improved the duplicate removal process.
- Implementing protective measures, including inmate screening, prophylactic programs, and outbreak-response protocols, could minimize the risk of infectious disease transmission in prisons.
- The results suggest the need for improved food safety protocols and stricter oversight of food preparation practices within correctional facilities.

Introduction

Specific conditions prevalent in penitentiary institutions can significantly increase the risk of infectious disease transmission among incarcerated populations. Several factors contribute to this heightened vulnerability in prison settings, including overcrowding, limited access to safe drinking water, delayed diagnosis, and poor ventilation.[1] Existing literature consistently reports elevated rates of infectious diseases within correctional facilities worldwide, including tuberculosis (TB), HIV (which can progress to AIDS), hepatitis C virus (HCV), and COVID-19.[2–10] Furthermore, the particular conditions inherent in prison environments may facilitate the transmission of foodborne illnesses and infections caused by ectoparasites such as mites.

Tuberculosis, caused by the bacterium *Mycobacterium tuberculosis*, is notably more prevalent in prisons than in the general population and remains one of the leading causes of death among incarcerated individuals.[1–3] Studies show that the prevalence of active TB is closely linked to the country in which the prison is located. In high-burden TB countries, the incidence of active TB is 2.5 times higher than in countries with a lower burden.[3] Tuberculosis typically presents as a pulmonary infection and is transmitted through inhalation of aerosolized droplets from an infected individual. Its constitutional symptoms are nonspecific and usually include a persistent cough, fever, weight loss, night sweats, and general malaise.[11]

According to data from the European Centre for Disease Prevention and Control (ECDC), presented in the Annual Epidemiological Reports (AERs), 30 countries within the European Union/European Economic Area (EU/EEA) reported a total of 36,179 TB cases in 2022 (8.0 per 100,000 population). Of these, 25.6% occurred in Romania (9,270 cases), while only 1 case was recorded in Liechtenstein.[12] Compared with previous years, a slight increase was noted. The number of reported drug-resistant TB cases in the EU/EEA also rose in 2022, which can be attributed to factors such as the resumption of regular testing after the COVID-19 pandemic, migration resulting from the war in Ukraine, and the expansion of targeted TB screening programs for high-risk populations.

Although the World Health Organization (WHO) has set a target of 85% for TB treatment success, the actual success rate reported in 2022 was only 64.0%, highlighting the need for significant improvement.[12] The *Status Report on Prison Health in the WHO European Region 2022*, which includes data from 16–28 member states, reported a TB prevalence of 0.46% among prison populations.[13]

A study published in 2021 found that TB cases among incarcerated individuals in Central and South America increased by 269% between 2000 and 2018. The prevalence of TB is disproportionately concentrated among people in prison, with 11% of all TB cases occurring within this population. In the past year, TB outbreaks have been reported in prisons in both Brazil and Venezuela.[14]

Scabies is an ectoparasitic infestation caused by *Sarcoptes scabiei* var. *hominis*, an obligate human parasite that completes its entire life cycle on the host. Female mites burrow into the stratum corneum of the skin to lay eggs, triggering an immune response that causes intense itching (pruritus) and erythematous rashes. The mites can survive for up to 24–36 h in environments with temperatures around 21°C and relative humidity levels between 40% and 80%, thus posing a risk of transmission during this period. Scabies is primarily spread through prolonged skin-to-skin contact with an infected individual. Symptoms usually appear 4–6 weeks after the initial infestation and typically include generalized itching, which worsens at night, along with a vesicular or papular rash.[15,16] In 2017, the WHO classified scabies as a Neglected Tropical Disease (NTD). However, epidemiological data on scabies remain limited in many countries, with most estimates based on a small number of non-representative surveys. Current estimates suggest a global point prevalence of 100–200 million scabies cases, with approx. 455 million new cases occurring each year. Scabies is estimated to cause around 3.8 million disability-adjusted life years (DALYs), making it one of the highest-burden NTDs.[17]

Salmonella is a genus of bacteria in the Enterobacteriaceae family and a major cause of gastroenteritis in humans. Transmission occurs primarily through ingestion of contaminated food or water, with poultry and poultry-derived products serving as key reservoirs. The severity of illness depends on the specific *Salmonella* serotype and the immune status of the host. The infection is classified into 2 main forms: typhoidal, which is characterized by persistent fever, headache, abdominal pain, anorexia, bradycardia, and either diarrhea or constipation; and non-typhoidal, which presents acutely with diarrhea, abdominal cramps, and fever.[18]

In 2022, 30 EU/EEA countries reported 66,721 cases of salmonellosis, of which 65,967 were laboratory confirmed. The highest numbers were recorded in the Czech Republic (7,563 cases) and Slovakia (3,669), while the lowest were reported in Portugal (412), Bulgaria (310), and Latvia (90). Although the number of cases increased compared with 2020, it remained lower than pre-COVID-19 levels.[19]

Shigellosis is an infection caused by bacteria of the *Shigella* genus, a group of facultative anaerobic, Gram-negative bacilli. Transmission occurs via the fecal–oral route, typically through direct person-to-person contact, ingestion of contaminated food or water, or, in some cases, through sexual contact, particularly among men who have sex with men. The hallmark clinical manifestation of shigellosis is dysentery, characterized by bloody, often mucous-laden diarrhea accompanied by abdominal cramps and vomiting.[20]

In 2022, 30 EU/EEA countries confirmed a total of 4,149 shigellosis cases. The highest numbers were reported in France (1,290 cases), Spain (598), and the Netherlands (419). Two countries (Iceland and Liechtenstein) reported no cases. Compared with 2020 and 2021, an increase in incidence was observed.[21]

Objectives

The aim of this study was to evaluate the incidence and epidemiological trends of TB, scabies, *Salmonella*, and *Shigella* infections among prisoners in Poland between 2002 and 2023, as well as to analyze the available literature addressing these diseases. The literature was sourced from publications indexed in Web of Science (WoS) and Scopus and published between 2015 and 2025. Additionally, this study sought to identify the most effective infection prevention strategies for implementation in penitentiary settings.

Materials and methods

This research consisted of 2 parts. First, a systematic review was conducted to analyze the existing literature on infectious diseases in prison populations, with a particular focus on TB, *Salmonella*, *Shigella*, and scabies. Second, findings from the authors' own empirical research are presented. Ethical review and approval were waived for this study because public databases were used.

Systematic review

Strategy of searching and criteria of material selection

The research protocol was prepared based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Fig. 1). The available literature on TB, *Salmonella*, *Shigella*, and scabies infections among inmates was analyzed. To gather the literature, 2 databases were searched: Scopus and WoS. The inclusion and exclusion criteria were as follows:

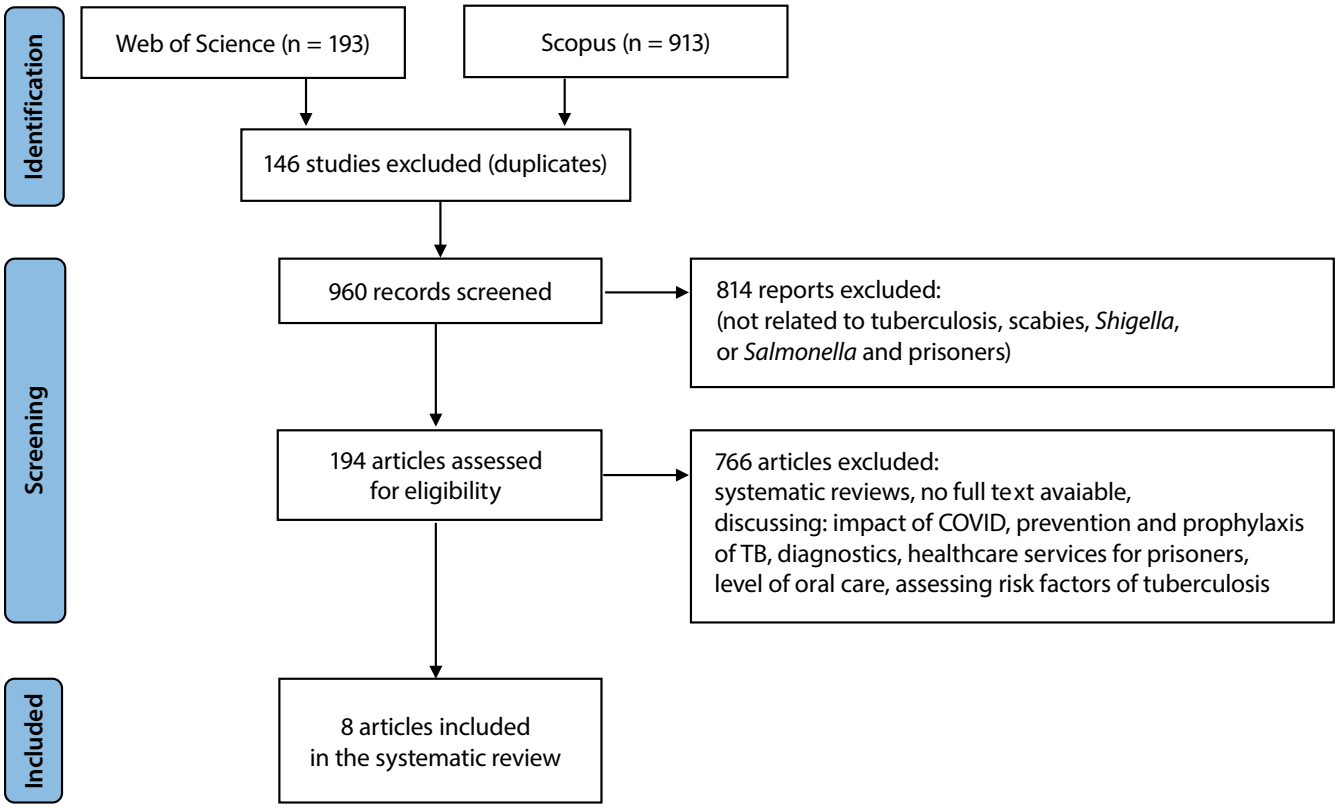


Fig. 1. Study selection process presented using a PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart

Inclusion criteria: original articles presenting TB (diagnosed using X-ray findings or epidemiological data from national institutions), *Salmonella*, *Shigella*, and scabies infections in prisoners; articles published in English between January 1, 2015 and April 9, 2025; full-text articles available; and studies assessed as satisfactory using the Newcastle–Ottawa Scale. Exclusion criteria: case reports, reviews, articles without full-text availability, and articles discussing the impact of COVID-19, prevention and prophylaxis of TB, diagnostics, healthcare services for prisoners, or risk factors for TB.

For database searching, the following keywords were selected: "prison" OR "inmates" OR "prisoners"; "infectious diseases"; "TB" OR "scabies" OR "*Shigella*" OR "*Salmonella*". Medical Subject Headings (MeSH) were used to identify appropriate search terms. Studies were screened by title and abstract according to the PICO (population, intervention, control, and outcome) criteria. The PICO question was as follows: Are prisoners at an increased risk of infection with TB, scabies, *Salmonella*, or *Shigella* compared with the general population?

The search was conducted on April 9, 2025. Article selection was performed independently by 2 calibrated researchers (A.T. and R.K.). Agreement between the reviewers was assessed using Cohen's kappa coefficient. The results were interpreted as follows: <0.00 – poor; 0.00–0.20 – slight; 0.21–0.40 – fair; 0.41–0.60 – moderate; 0.61–0.80 – substantial; and 0.81–1.00 – almost perfect. The agreement between the reviewers was 0.442.

Duplicates exclusion

To eliminate duplicates, a Python script was used. The script combined and processed 2 bibliographic datasets from WoS and Scopus to generate a unified list of medical publications. It began by reading Microsoft Excel 2013 (Microsoft Corp., Redmond, USA) and CSV files using the Pandas library, extracting relevant columns such as authors, title, source, and abstract. The script then standardized column names across both datasets and merged them into a single DataFrame. To remove duplicate entries, it created a "clean_key" column by stripping and normalizing article titles. After duplicate removal, source names were formatted in title case and the DataFrame index was reset. The final dataset was used to generate a Word document using the python-docx library. Each publication entry was formatted with bold labels and listed as a numbered item. The final document included the title, authors, source, and abstract for each article. This approach ensured consistent formatting and removed redundancies across multiple bibliographic sources (Fig. 2).

Substantive analysis

The following information was extracted from the selected manuscripts: year of the study, country in which it was conducted, characteristics of the study participants, and information on the prevalence of selected infectious diseases among inmates. Data were extracted by 1 researcher (A.T.). The extracted information was checked by another co-author (M.T.) to minimize the risk of bias.

Quality assessment

The reliability of the studies was assessed using the Newcastle–Ottawa Quality Assessment Scale for cross-sectional studies. The following domains were assessed: selection (SD), comparability (CD), and outcome (OD). The maximum score was 10 points: 5 points for the 1st category, 2 points for the 2nd category, and 3 points for the outcome domain. Interpretation of the Newcastle–Ottawa Quality Assessment Scale was as follows:

– 9–10 points: very good quality;
– 7–8 points: good quality;
– 5–6 points: satisfactory quality;
– 0–4 points: unsatisfactory quality.

This part of the research was conducted independently by 2 authors (A.T. and R.K.). Any discrepancies were resolved by the 3rd author (M.R.).

Analysis of epidemiological data

The research methodology was based on the possibility of obtaining anonymized data concerning inmates in all prisons across the country diagnosed with TB, scabies, *Salmonella*, and *Shigella* infections between 2002 and 2023. Access to these sensitive data was possible due to the obligation of the Prison Service (PS) to provide health-related information about inmates to Statistics Poland, in accordance with the provisions of the Act of June 29, 1995, on Public Statistics,[22] as well as the annually updated Program of Statistical Research of Public Statistics (PSRPS). According to the Public Statistics Act, Article 13: "Respondents, including entities in the public finance sector, are obliged to provide statistical data free of charge for the purposes of public statistics." Health data of prisoners, transmitted by the PS to statistical offices, are collected and forwarded on the basis of medical reports and healthcare registry systems within correctional facilities. These data are gathered through internal registry systems that maintain detailed medical documentation of inmates. Subsequently, the PS submits this information to statistical offices, such as Statistics Poland, for the purposes of analysis, reporting, and use in social and health research. For this study, publicly available data from the PS official website (https://sw.gov.pl/strona/Statystyka) were used. The information contained there is collected through internal recording systems based on medical reports related to inmate healthcare.

Due to the retrospective and noninvasive nature of the study, and because all data were anonymous, ethics committee approval was not required. Studies exempt

```python
import pandas as pd
from docx import Document

# Load publication data from Web of Science (Excel format)
publ1 = pd.read_excel(r"C:\medical_bases\web_of_science_medical_base.xls")

# Load publication data from Scopus (CSV format)
publ2 = pd.read_csv(r"C:\medical_bases\scopus_medical_base.csv")

# Select relevant columns from each source
publ1_extract = publ1[['Authors', "Article Title", "Source Title", "Abstract"]]
publ2_extract = publ2[['Authors', 'Title', 'Source title', 'Abstract']]

# Standardize column names for consistency
publ1_extract.columns = ['Authors', 'Title', 'Source', 'Abstract']
publ2_extract.columns = ['Authors', 'Title', 'Source', 'Abstract']

# Combine both datasets into a single DataFrame
publ_all = pd.concat([publ1_extract, publ2_extract], axis=0)

# Create a normalized key based on the title to identify duplicates
publ_all['clean_key'] = (
    publ_all['Title']
    .str.strip()                                  # Remove leading/trailing whitespace
    .str.lower()                                  # Convert to lowercase
    .str.replace(r"[^a-z0-9]", "", regex=True))  # Remove non-alphanumeric characters

# Drop duplicate records based on the cleaned title
publ_all_clean = publ_all.drop_duplicates(subset='clean_key', keep="first")

# Remove the helper column used for deduplication
publ_all_clean = publ_all_clean.drop(columns='clean_key')

# Format the source title into title case for consistency
publ_all_clean['Source'] = publ_all_clean['Source'].str.title()

# Reset the index to start from 1 for cleaner document formatting
publ_all_clean = publ_all_clean.reset_index(drop=True)
publ_all_clean.index = publ_all_clean.index + 1

# Initialize a Word document
doc = Document()

# Write each publication entry as a numbered paragraph with formatted fields
for i, row in enumerate(publ_all_clean.itertuples(index=False), start=1):
    p = doc.add_paragraph(style='List Number')  # Create numbered list item

    # Add formatted publication details
    p.add_run("Title: ").bold = True
    p.add_run(f"{row.Title}\n")

    p.add_run("Authors: ").bold = True
    p.add_run(f"{row.Authors}\n")

    p.add_run("Source: ").bold = True
    p.add_run(f"{row.Source}\n")

    p.add_run("Abstract: ").bold = True
    p.add_run(f"{row.Abstract}")

# Save the resulting document to disk
doc.save("Identified_articles.docx")
```

Fig. 2. Python script used for duplicate elimination

from ethical approval include most educational research, case reports on 1–3 patients (without hypothesis testing), studies involving no risk to participants, studies involving information available in the public domain, analyses of open-source or anonymized datasets obtained from other researchers with appropriate informed consent collected during primary data collection, and research evaluating public health programs or government schemes.[23]

Statistical analyses

The Shapiro–Wilk test was conducted to evaluate whether the distribution of a given variable conformed to a normal distribution. The test results indicated a significant deviation from normality ($p < 0.050$). Consequently, it was deemed appropriate to utilize Spearman's rank correlation coefficient instead of Pearson's correlation coefficient, as the former does not require the assumption of normality in the data distribution. Using Spearman's rho correlation coefficient, the association between variables that did not follow a normal distribution was assessed. The p-values were determined using an asymptotic approximation based on the t-test. The p-values were additionally adjusted (p_{adj}) using the Holm method when multiple comparisons were assessed.

The growth rate of the time-series data was quantified using the compound annual growth rate (CAGR) parameter, calculated with the following formula (https://www.calkoo.com/en/cagr-calculator):

$$CAGR = (EV/BV)^{(1/n)} - 1,$$

where: EV – the value of the studied parameter in the final year; BV – the initial value in the first year; n – the number of years under observation. Analyses were conducted using R v. 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria) on Windows 10 Pro 64-bit (build 19045; Microsoft Corp.), using the packages sjPlot v. 2.8.15 (https://strengejacke.github.io/sjPlot), report v. 0.5.7,[24] ggstatsplot v. 0.12,[25] and ggplot2 v. 3.4.4.[26] Microsoft Excel 2013 and Statistica v. 11.0 (StatSoft Inc., Tulsa, USA) were used to carry out the analyses.

Results

Systematic review

Quality assessment

The Newcastle–Ottawa Scale was used to assess the quality of the 8 articles included in the review. This part of the research was conducted independently by 2 authors (A.T. and R.K.). Any discrepancies were resolved by the 3rd author (M.R.) (Table 1). Only 1 article was assessed as satisfactory, 2 as good, and 5 as very good.

Study characteristics

Two databases were searched: Scopus and WoS. As a result, 1,106 articles were identified (Fig. 1). After removing duplicates, 960 articles were obtained for the first screening. In the next step, the articles were analyzed with regard to the inclusion criteria, which resulted in the exclusion of 814 articles that did not address the prevalence of the selected infectious diseases (TB, *Salmonella*, *Shigella*, or scabies). Subsequently, 766 articles were excluded for the following reasons: being systematic reviews, unavailability of the full text, or focus on topics such as the impact of COVID-19, prevention and prophylaxis of TB, diagnostics, healthcare services for prisoners, level of oral care, or assessment of TB risk factors.

As a result, 8 articles were included in the review (Table 2). Each article described the situation in a different country (USA, Côte d'Ivoire, North Macedonia, Iran, Gauteng, Limpopo, Mpumalanga and North West Provinces in South Africa, China, Poland, and Ghana). One article was published in 2016, 1 in 2017, 2 in 2018, 2 in 2019, 1 in 2020, and 1 in 2024. Six of the articles discuss TB infections,[27–32] while 2 address scabies infections in prisons.[33,34]

None of the articles included in the review addressed the problem of *Salmonella* or *Shigella* infections. During article screening, the authors identified 1 article describing *Salmonella*; however, because it was a case report, it was excluded from the review.

Table 1. Results of the quality assessment using the Newcastle–Ottawa Scale

Reference	Selection				Comparability	Outcome		Quality
	A	B	C	D		A	B	
27	1	1	0	1	2	2	0	good
28	1	1	1	1	2	2	1	very good
29	1	1	1	1	2	2	0	good
30	1	1	1	1	1	0	0	satisfactory
31	1	1	1	1	2	2	1	very good
32	1	1	1	1	2	2	1	very good
33	1	1	1	1	2	2	1	very good
34	1	1	1	1	2	2	1	very good

Table 2. Articles included in the review

Authors, title, and year of publication	Country	Participants/material	Statistical analysis
Lambert et al., Tuberculosis in jails and prisons: United States, 2002–2013; 2016[27]	USA	2002–2013 National Tuberculosis Surveillance System case reports	Statistical Analysis System v. 9.3 (SAS Institute Inc, Cary, USA)
Séri et al., Prevalence of pulmonary tuberculosis among prison inmates: A cross-sectional survey at the Correctional and Detention Facility of Abidjan, Côte d'Ivoire; 2017[28]	Côte d'Ivoire	943 inmates screened	univariate and multivariate logistic regression, SAS v. 9.3
Ilievska-Poposka et al., Tuberculosis in the prisons in the Republic of Macedonia, 2008–2017; 2018[29]	North Macedonia	data from Central Register for TB	–
Seyedalinaghi et al., Comparing tuberculosis incidence in a prison with the society, Tehran, Iran; 2018[30]	Iran	5865 prisoners 4,503,516 people of the society	–
Jordan et al., Prevalence and risk factors of tuberculosis disease in South African correctional facilities in 2015; 2019[31]	Gauteng, Limpopo, Mpumalanga, and North West Provinces in South Africa	57,660 inmates	bivariate logistic regression, multivariate regression model
Tong et al., Epidemic situation of tuberculosis in prisons in the central region of China; 2019[32]	China	3,459 prisoners examined and compared to epidemic situation in the province and general situation in China	χ^2 test SPSS v. 13.0 (SPSS Inc., Chicago, USA)
Bartosik et al., Scabies and pediculosis in penitentiary institutions in Poland: A study of ectoparasites in confinement conditions; 2020[33]	Poland	data collected in prisons provided by the Central Board of Prison Service	Spearman's rank correlation coefficient, Pearson's correlation coefficient
Amoako et al., Burden of scabies in Ghanaian penitentiary; 2024[34]	Ghana	559 prisoners	frequencies, proportions, median, interquartile range, χ^2 test of association, Fisher's exact test, Mann–Whitney U test

TB – tuberculosis.

Tuberculosis in prisons

The problem of TB in penitentiaries is widely discussed in the available literature. In 4 articles, the authors performed X-rays to confirm the TB diagnosis.[28,30–32] Séri et al.[28] confirmed the diagnosis through medical examination and interview, followed by chest X-rays and sputum collection. Similar methodology was used by authors from Iran, South Africa, and China.[30–32] Researchers from North Macedonia and the USA decided to analyze data provided by national institutions. Lambert et al.[27] analyzed data obtained from the National Tuberculosis Surveillance System. Authors from North Macedonia gathered information from the Central Register for TB within the Institute for Lung Diseases.[29] Table 3 presents the results of the research described in the studies included in the review.

Scabies in penitentiary institutions

Bartosik et al.[33] analyzed data provided by the Central Board of the Prison Service and the District Inspectorate of the Prison Service in Lublin. They also described the problem of pediculosis in prisons. The 2nd article included in the review and discussing the problem of scabies was authored by researchers from Ghana.[34] They conducted a cross-sectional study, collecting information from medical histories and demographic data using the Research Electronic Data Capture (REDCap) questionnaire. Subsequently, skin examination was performed and scabies was diagnosed according to the criteria of the International Alliance for the Control of Scabies (ICAS).[34] The results of these 2 studies are presented in Table 4.

Epidemiological data analysis

Analysis of X-ray screenings and tuberculosis detection in Polish prisons (2002–2023)

Over the 22-year period, a discernible trend emerges from the analysis of X-ray screenings and TB case detection (Fig. 3). Initially, a slight increase in the number of X-rays performed

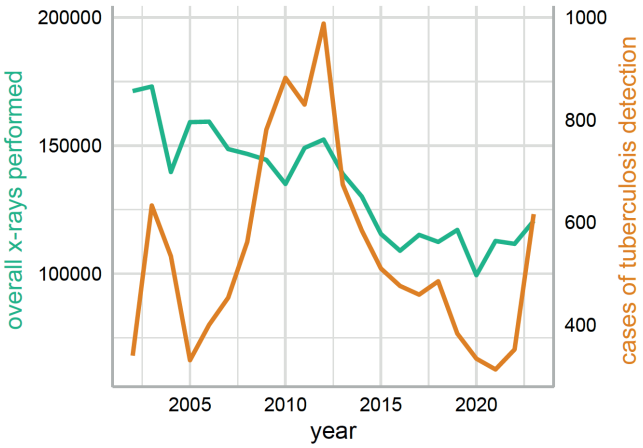


Fig. 3. Annual trends in X-ray screenings and tuberculosis cases in Polish detention wards and penitentiaries (2002–2023)

Table 3. Data regarding prevalence of TB in prisoners presented in the articles included in the review

Reference	Inmate characteristics	Results	Conclusions
Lambert et al.[27]	495 female patients diagnosed with TB 5,084 male patients diagnosed with TB	– 45% of prisoners diagnosed with TB were U.S.-born; – 9% were female; – median estimated annual TB incidence rates were: 29 cases per 100,000 in local jails, 8 per 100,000 in state prisons, 25 per 100,000 in federal ones; – between 2009 and 2013 TB screenings were leading to 47% of diagnoses among female inmates and 42% among male inmates	Systematic screening and treatment of TB are essential for TB prevention and control.
Séri et al.[28]	943 inmates (median age of 31): 871 from the men's building (median age of 31) 62 from the women's building (median age of 29) 10 from the infirmary (median age of 32)	– 88 inmates (9.3%) met the TB case definition: – 19 (2%) – confirmed TB, – 40 (4.2%) – probable TB, – 29 (3.1%) – possible TB; – there were 19 TB strains isolated: 10 (53%) were drug-resistant, including 7 (37%) with multiresistance; – factors associated with confirmed TB were: age ≥30 years (OR = 3.8; 95% CI: 1.1–13.3), prolonged cough (OR = 3.6; 95% CI: 1.3–9.5), and fever (OR = 2.7; 95% CI: 1.0–7.5)	– pulmonary TB is 10–44 times as frequent as in the general population; – there is a need of implementation of the annuals in-cell TB screening campaigns
Ilievska-Poposka et al.[29]	58 prisoners with TB: 57 (98.2%) were male 1 (1.72%) case was female; the youngest prisoner with TB was 20, the oldest – 59 years old	– the incidence rate of TB is higher than 100/100,000 population, or several times higher than in general population (except 2012 and 2016); – the majority of inmates diagnosed with TB were young men with risk factors for TB or TB before incarceration; – the most of the patients diagnosed with TB in prison were new cases (54); – 82.75% of TB cases were successfully treated, there were no cases of multidrug-resistant TB	– TB control in prisons is good; the treatment outcomes are satisfactory; – prisons in North Macedonia provide conditions for TB transmission
Seyedalinaghi et al.[30]	9 prisoners (out of 5,865) age range 23–49 years diagnosed with TB 2,263 people from the society out of 4,503,516 with the same age diagnosed with TB	– out of 100,000 individuals, 153.5 patients were diagnosed in prison and 5.02 in the society; – the incidence of TB in the prison was significantly higher than in the society ($p < 0.0001$) and the ratio of TB incidence in prison to the society was 30.6 (95% CI: 16–58)	– the incidence of TB in prisons is significantly higher when compared to the general population; – there is a need for preventive measures in prisons, such as intensifying case findings in prisons
Jordan et al.[31]	the overall inmate population was 57,660; 31,668 were screened for TB, 175 reported current TB treatment; 837 inmates with TB	– the estimated total prevalence rate of TB (2,653/100,000) and estimated prevalence rate of screening-identified TB (2,116/10,000) indicated higher TB burden in prisons than in the global general population (174/100,000) or in South Africa in 2015 (520/100,000); – previous TB was associated with increased odds of screening-identified TB in HIV-positive inmates (OR = 4.3, 95% CI: 2.5–7.3); – in HIV-negative inmates, previous TB (adjusted OR = 4.9, 95% CI: 1.7–14.1) and self-reported symptoms vs none were independently associated with increased odds of screening-identified TB	There is a need for routine TB screening, including computer-assisted digital chest X-ray to identify and refer prisoners with active TB.
Tong et al.[32]	3,459 prisoners (1,133 men, 2,326 women) underwent chest X-ray examination; 40 with active TB (18 men, 22 women)	– the active TB prevalence in prison was significantly higher than that of the province and in Chinese general population – 1156/100,000 ($p < 0.01$); – TB prevalence in men's prison (1,589/100,000) was higher than in women's (946/100,000); – the risk of TB infection in women's prison was much higher than in men's when compared with the TB prevalence from the province (woman: OR = 2.37, 95% CI: 1.34, men: OR = 1.53, 95% CI: 0.90–2.60) and the China's general population (women: OR = 3.30, 95% CI: 2.15–5.09; men: OR = 2.06, 95% CI: 1.29–3.30)	In case of the severe epidemic situation of TB in prisons, integrating medical resources to establish a consummate and effective management system is necessary.

TB – tuberculosis; OR – odds ratio; 95% CI – 95% confidence interval.

was observed, peaking at 173,047 in 2003, followed by a general downward trajectory to a low of 99,400 in 2020.

Parallel to the changing frequency of X-ray use (CAGR = –1.9%; 95% confidence interval (95% CI): –2.9 to –1.9), the number of TB cases detected showed substantial fluctuation, not strictly following the pattern of X-ray screenings. A significant rise in TB cases was noted from the early to mid-2010s, peaking in 2012 with 988 cases. Subsequently, a decrease in TB cases after 2012 was observed, reaching a nadir in 2021, followed by a resurgence to 616 cases in 2023.

Table 4. Data regarding scabies prevalence in inmates presented in the articles included in the review

Reference	Inmate characteristics	Results	Conclusions
Bartosik et al.[33]	number of scabies cases (2001–2015): 28,943 (prevalence: 2.3%)	– the lowest number of scabies cases was noted in 2007–2008 (1,115 and 1,103), the highest in 2001–2002 (3,072 and 3,071); – after 2003, number of scabies cases decreased till 2010, then an increase was observed; – the median number of scabies cases between 2001–2015 was 2029 (standard deviation = 627); – the highest prevalence of scabies was observed in 2001, 2002 and 2014, the lowest percentage in 2003–2010; – the number of prisoners does not correlate with the increase in the scabies cases (the greater the number of prisoners, the lower the number of infection cases was; Spearman's rho = –0.718; p = 0.003)	– prisoners should be carefully monitored for the presence of scabies; – education of inmates could be a promising tool in disease prevention
Amoako et al.[34]	559 prisoners (540 male, 19 female) median age was 36 (33–40) years	– 368 (65.8%) diagnosed with scabies, no female inmate; – scabies was mild in 63.3% and moderate in 30.7% (lesions found on the hands, wrists, fingers, and finger webs); – 17% of prisoners with scabies were also diagnosed with impetigo	– very high prevalence of scabies among inmates; – mass drug administration with ivermectin and health education are needed to reduce the number of scabies infections in inmates

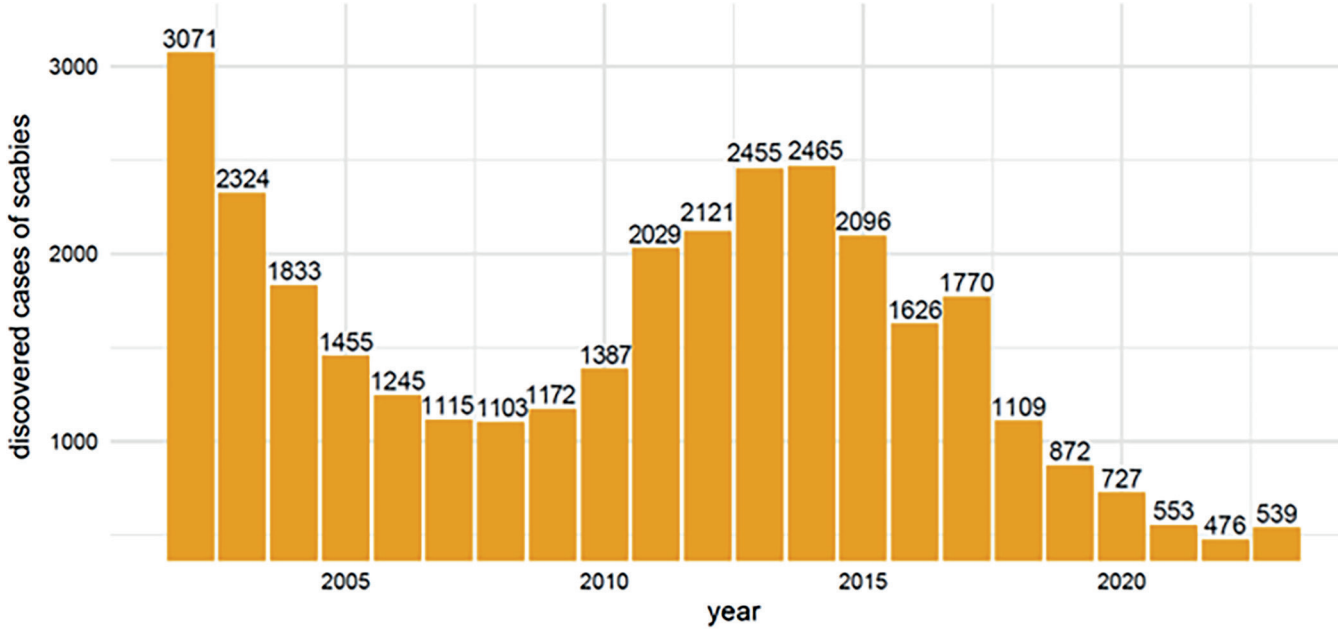


Fig. 4. Annual trends in reported scabies cases in Polish detention wards and penitentiaries (2002–2023)

Trends and compound annual growth rate analysis of scabies cases in Polish prisons (2002–2023)

The data in Fig. 4 represent the number of scabies cases reported annually from 2002 to 2023. Initially, there was a consistent decrease in cases from 3,071 in 2002 to 1,103 in 2008. However, from 2009 onwards, there was a noticeable reversal in this trend, with cases increasing and peaking in 2014 at 2,465 cases. After reaching the peak in 2014, the number of cases began to decline again. By 2021, the number of cases had dropped significantly to 553, reaching a new low in 2022 with 476 cases, before a slight increase to 539 cases in 2023. This negative CAGR of approx. –2.4% (95% CI: –3.6 to –1.3) annually over the 21-year period indicates a general decline in scabies cases despite the observed fluctuations. This overall decline suggests that the measures implemented over the years have been largely effective, albeit with periods of challenge requiring renewed focus and possible adaptation of strategies.

The analysis of trends in penitentiary capacity and inmate population in Poland from 2002 to 2023, based on the data presented in Fig. 5, reveals the dynamics of the prison system over 2 decades. Overall, the data from 2002 to 2023 reveal a correctional system that has undergone significant changes aimed at effectively balancing capacity with inmate populations.

Initially, both capacity and the number of inmates showed an upward trend. The capacity of penitentiaries increased steadily from 69,083 in 2002 to a peak of 87,742 in 2014, reflecting a systematic expansion of facilities, which coincided with an increase in the inmate population, reaching a peak of 89,995 in 2007 (Fig. 5).

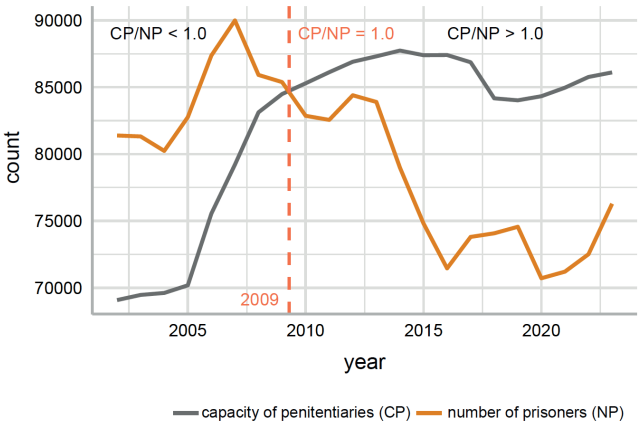


Fig. 5. Dynamics of penitentiary capacity and prisoner numbers in Poland from 2002 to 2023

Annual trends in diagnostic examinations for *Salmonella*/*Shigella*: volume, positivity rates, and comparative analysis in Polish prisons (2002–2023)

Analyzing the data on diagnostic tests for *Salmonella*/*Shigella* in Polish inmates from 2002 to 2023 reveals interesting trends in both the number of tests conducted and the positivity rates for each disease.

For *Salmonella*/*Shigella*, the number of tests initially hovered around 5,000 in the early years, gradually increasing to a peak of 8,876 in 2020. Following this peak, there was a noticeable decline, with 2023 recording 5,204 tests (CAGR = −0.77%; 95% CI: −2.50 to 1.10). The positivity rates for these tests were generally low, often below 1%, except for certain years such as 2013 and 2014, when they increased to 2.48% and 1.72%, respectively (Fig. 6).

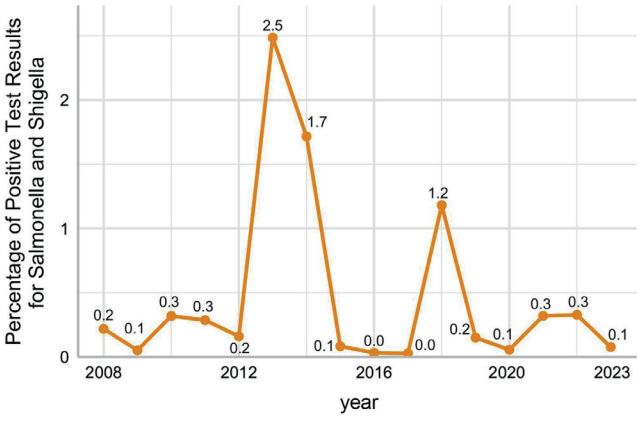


Fig. 6. Annual trends in positive results for *Salmonella*/*Shigella* testing among Polish inmates (2002–2023)

Discussion

The effective implementation of this research project, which enabled the collection of selected anonymized statistical data from specific study cohorts – particularly inmates – was made possible by the robust operation of the statutory obligation requiring the PS to transmit health-related data concerning inmates to Statistics Poland. The PS is legally authorized to process personal and sensitive data, including inmates' health information, without explicit consent when such processing is necessary to fulfill statutory duties, maintain order and security within penitentiary institutions, and provide appropriate living and healthcare conditions. All data processing is conducted in strict compliance with data protection regulations, with full respect for individual dignity and the right to privacy. Personal data may be processed and disseminated in a manner that precludes identification of individuals, particularly for statistical and research purposes.[35]

Although prison systems vary globally, common challenges are observed across countries – overcrowding, poor inmate health, inactivity, violence and abuse, substance use disorders, and limited access to healthcare – each representing major public health concerns.[36] In 2023, the global occupancy rate in penitentiary centers varied significantly: 616.9% in the Republic of the Congo (Congo-Brazzaville), 302% in Haiti, 231.4% in Peru, 176.5% in Kenya, 101.5% in Hungary, 95.6% in the USA, 77.6% in Germany, 67.0% in Russia, 56.6% in Japan, 31.0% in Monaco, and 14% in Nauru. In Poland, the occupancy rate was reported at 87.2%.[37,38]

Despite an increase in prison capacity, the prisoner-to-capacity ratio in Poland exceeded 1.0 until 2009, indicating persistent overcrowding. From 2010 onward, this ratio dropped below 1.0, suggesting either a reduction in inmate numbers or effective expansion of prison infrastructure. Notably, after 2014, both the number of prisoners and available capacity declined – although the inmate population fell at a faster rate. By 2016, the ratio had reached its lowest point (0.82) and remained below 0.9 through 2023. These trends may reflect successful reforms such as sentencing alternatives, non-custodial penalties, and recidivism reduction programs. Overcrowding is a well-established factor contributing to the spread of infectious and parasitic diseases, including TB, scabies, *Salmonella*/*Shigella* infections, and HIV. Our findings show that increases in TB and scabies cases are closely correlated with prison population density. The highest number of TB cases was recorded in 2012 (988 cases), yet this did not coincide with the peak in X-ray screenings, indicating that other factors – such as overcrowding or more virulent TB strains – may have contributed. A decline followed, reaching a low point in 2021, before resurgence in 2023. The weak correlation between case rates and X-ray screening volume (rho = 0.28, p = 0.203, 95% CI: −0.16 to 0.63) suggests that screening alone is insufficient to control TB and should be part of a broader health strategy.

The decrease in screening rates and the complex trend in TB incidence (CAGR = 2.5%, 95% CI: 0.6–4.4) reflect the dynamic nature of healthcare challenges in prisons. This underscores the need for ongoing adaptation – through advanced diagnostics, improved control measures, and better healthcare access for inmates. Notably,

the significant drop in screenings in 2020 likely resulted from the COVID-19 pandemic disrupting routine care. Tuberculosis remains a strong indicator of overall infection control performance, and its persistence raises broader concerns about institutional preparedness for managing infectious disease outbreaks, including SARS-CoV-2.[39]

Our findings align with those of Placeres et al.,[3] whose meta-analysis confirmed TB in prisons as a global public health and social issue, advocating rapid assessments and precision interventions in correctional settings. Several other studies confirm the elevated TB burden in prisons,[27–32] including work by Séri et al.,[28] Ilievska-Poposka et al.,[29] Seyedalinaghi et al.,[30] Jordan et al.,[31] and Tong et al.,[32] all of whom emphasized the urgent need for effective TB management systems in penitentiaries.

Implementing protective measures, including screening of inmates, prophylactic programs, and protocols in case of an outbreak, could minimize the risk of infectious disease transmission in prisons. Beaudry et al.[40] conducted a systematic review on managing outbreaks of highly contagious diseases in prisons. Out of 28 studies included in the review, 14 agreed that screening was crucial to reduce the causes of the outbreak. Kending et al.[41] reviewed the challenges and strategic opportunities for infection prevention and control (IPC) efforts in U.S. jails and prisons. They recommended several steps that could improve IPC: training personnel to manage IPC programs, improving infectious disease surveillance, preparing augmented evidence-based guidance particularly for correctional facilities, strengthening external stakeholder engagement (vaccination programs, screening, outbreak management), and developing and funding a national research agenda to evaluate IPC implementation. In the USA, scientific reports have identified a correlation between the rise of AIDS and the resurgence of TB in prisons, particularly since the beginning of the HIV/AIDS epidemic in 1981.[42]

Scabies remains a public health challenge both in Poland and globally.[43–46] Korycińska et al.[43] found that scabies prevalence was highest in urban and rural communities, correlating with reduced healthcare access and lower health literacy. Adolescents (aged 10–19 years) were most affected, likely because of increased close contact. The infection also correlated with socioeconomic status, including residence in areas with high unemployment.

A systematic review by Delie et al.[16] linked scabies in prisons to poor hygiene, overcrowding, shared beds and clothing, and shorter sentence lengths. Similar results were presented by Bogino et al.,[5] who found strong associations with inadequate hygiene, shared clothing, lack of soap, and poor ventilation.

In our analysis, scabies incidence declined from 2002 to 2008, rose to a peak in 2014, and then declined again. These fluctuations may reflect varying levels of disease control, changes in reporting practices, or other institutional factors. The likely driver of the mid-period increase was prison population growth. However, Bartosik et al.[33] found no clear correlation between inmate numbers and scabies cases in Polish prisons (2001–2015), with higher inmate counts sometimes coinciding with fewer cases.

Only 2 studies in our review discussed scabies prevalence in Poland and Ghana.[33,34] Between 2001 and 2015, Polish data showed a mean infection rate of 2.3% (min 1.2% in 2007; max 3.9% in 2001), while at Kumasi Central Prison in Ghana, scabies prevalence reached 65.8%. This stark contrast highlights the critical importance of control measures, hygiene, and health education.

The literature also documents outbreaks of *Shigella* and *Salmonella* in correctional facilities.[6–8] Ranjabar et al.[6] described a *Shigella flexneri* outbreak in Isfahan, Iran, linked to raw vegetable consumption. Gicquelais et al.[7] emphasized strict food safety practices after a *Salmonella* outbreak in 2 Arkansas prisons caused by contaminated eggs. Mardu et al.[8] advocated food-handler education as a means of outbreak prevention.

In our study, the highest positive test rates for *Shigella* and *Salmonella* were recorded in 2013 (2.48%) and 2014 (1.72%), respectively. This could suggest increased awareness and testing during that period or reflect actual outbreaks mitigated by staff training and better food safety practices.

Insfran-Rivarola et al.[47] in their systematic review analyzed the effects of food safety and hygiene training on food handlers. They included 31 studies in the review and classified them into 3 categories according to the main outcome. Sixteen articles reported changes in food safety practices, with interventions increasing adherence to food safety practices. The authors agreed that, for the training to be more successful, it is important to combine theory and practice, as active participation in training has been shown to be more effective. Moreover, they noted that such programs should be implemented regularly to prevent knowledge attrition.

Dadu et al.[48] reported a decline in TB incidence in European prisons (2014–2018), including in Russia. This trend likely reflects WHO-backed national TB control strategies launched in 2014. A similar pattern was observed in Poland, where TB incidence declined between 2014 and 2017. In Poland, prison healthcare is governed by the Regulation of the Minister of Justice, published on June 14, 2012 (https://eli.gov.pl/eli/DU/2012/738/ogl), with subsequent amendments likely contributing to improved TB control.

The use of a Python script was the authors' own approach, which enhanced the selection process during the article-screening stage. This method may be useful for other authors preparing a systematic review.

Limitations of the study

Several limitations must be acknowledged. First, we analyzed only secondary data provided by authorities; no direct patient examination was conducted. We do not have information on the basis on which the diagnoses were established or whether they were validated through laboratory testing. Second, we lack information on whether data collectors

were trained and standardized, which limits methodological comparability. Third, data on scabies were incomplete, and no data on *Salmonella* or *Shigella* infections were available before 2008, restricting trend analysis to more recent years.

Conclusions

Overcrowding in prisons, observed from 2002 to 2023, has contributed to the transmission of infectious and parasitic diseases, particularly scabies, as well as TB and *Salmonella*/*Shigella* infections. A marked increase in TB cases occurred from the early to mid-2010s, peaking in 2012. This rise does not correspond to the highest number of chest X-rays performed. The elevated number of positive test results for *Salmonella*/*Shigella* infections paralleled the growth of the prison population. This underscores the necessity for improved food safety protocols and more stringent oversight of food preparation practices within correctional facilities, with a particular emphasis on preventing gastrointestinal diseases and enhancing hygiene standards. To mitigate disease transmission within the prison environment, it is essential to implement rapid health assessments that enable targeted and precise interventions. Incarceration restricts personal liberties; however, this should not occur at the expense of reducing the minimum standard of healthcare, which should remain at least equivalent to that available in the general population. The available literature emphasizes the higher occurrence of TB compared with the general population and the need for implementation of effective systems to control TB and scabies in correctional settings.

The use of a Python script during the systematic review process (at the duplicate-removal stage) improved the process, saved time, and eliminated errors that may result from human factors. To minimize the risk of infectious disease transmission in prisons, several protective measures should be implemented. These include systematic screening of newly admitted inmates, introduction of prophylactic programs involving periodic screening, and development of clear response protocols in the event of an outbreak. Effective control of infectious diseases within correctional facilities requires coordinated efforts and collaboration among all relevant authorities.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15806506. The package contains the following files:

The data presented in the supplementary tables and figures show the number of cases of the studied diseases among prisoners in Poland in individual years. Supplementary Table 1 and Supplementary Fig. 1 present the number of TB cases among inmates in Poland from 2002 to 2023.

A significant rise in TB cases was observed from the early to mid-2010s, peaking in 2012 with 988 cases. This was followed by a decline in TB incidence after 2012, reaching its lowest point in 2021, and then rising again to 616 cases in 2023. Supplementary Table 2 and Supplementary Fig. 2 present the number of scabies cases. Initially, a consistent decline was observed from 3,071 cases in 2002 to a low of 1,103 cases in 2008. However, from 2009 onwards, this trend reversed, with a steady increase peaking in 2014 at 2,465 cases.

After reaching this peak in 2014, the number of cases began to decline again. Supplementary Table 3 and Supplementary Fig. 3 present data on the number of inmates diagnosed with *Salmonella* or *Shigella* infections. The highest number of cases was recorded in 2013, with a total of 121 cases.

Data Availability Statement

The data that support the findings of this study are available in Zenodo at https://doi.org/10.5281/zenodo.15852790.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Rafał Korkosz https://orcid.org/0000-0001-9608-1949
Agata Trzcionka https://orcid.org/0000-0003-4884-909X
Robert Deręgowski https://orcid.org/0009-0000-3696-116X
Maksymilian Kiełbratowski https://orcid.org/0009-0003-1399-7719
Anna Kuśka-Kiełbratowska https://orcid.org/0000-0001-8708-470X
Mansur Rahnama- Hezavah https://orcid.org/0000-0002-5622-7330
Marta Tanasiewicz https://orcid.org/0000-0002-8700-7381

References

1. Enggist S, Møller L, Galea G, Udesen C, eds. *Prisons and Health*. Copenhagen, Denmark: World Health Organization (WHO), WHO Regional Office for Europe; 2014. ISBN:978-92-890-5059-3.
2. Busatto C, Mespaque J, Schwarzbold P, et al. Tuberculosis in prison inmates in Southern Brazil: Investigating the epidemiological and operational indicators. *Rev Soc Bras Med Trop*. 2022;55:e0052-2022. doi:10.1590/0037-8682-0052-2022
3. Placeres AF, De Almeida Soares D, Delpino FM, et al. Epidemiology of TB in prisoners: A metanalysis of the prevalence of active and latent TB. *BMC Infect Dis*. 2023;23(1):20. doi:10.1186/s12879-022-07961-8
4. Tanaka TSO, Cesar GA, Rezende GRD, et al. Molecular epidemiology of HIV-1 among prisoners in Central Brazil and evidence of transmission clusters. *Viruses*. 2022;14(8):1660. doi:10.3390/v14081660
5. Bogino EA, Woldegeorgis BZ, Wondewosen L, et al. Scabies prevalence and its associated factors among prisoners in southern Ethiopia: An institution-based analytical cross-sectional study. *PLoS Negl Trop Dis*. 2023;17(12):e0011826. doi:10.1371/journal.pntd.0011826
6. Ranjbar R, Hosseini MJ, Kaffashian AR, Farshad S. An outbreak of shigellosis due to *Shigella flexneri* serotype 3a in a prison in Iran. *Arch Iran Med*. 2010;13(5):413–416. PMID:20804308.
7. Gicquelais RE, Morris JF, Matthews S, et al. Multiple-serotype salmonella outbreaks in two state prison: Arkansas, August 2012. *MMWR Morb Mortal Wkly Rep*. 2014;63(8):169–173. PMID:24572612. PMCID:PMC4584523.
8. Mardu F, Negash H, Legese H, et al. Assessment of knowledge, practice, and status of food handlers toward *Salmonella*, *Shigella*, and intestinal parasites: A cross-sectional study in Tigrai prison centers, Ethiopia. *PLoS One*. 2020;15(11):e0241145. doi:10.1371/journal.pone.0241145
9. Korkosz R, Trzcionka A, Mączkowiak D, et al. Dental treatment needs of male inmates in relation to the analysis of medical databases. *J Clin Med*. 2024;13(3):858. doi:10.3390/jcm13030858

10. Korkosz R, Trzcionka A, Hildebrandt T, et al. Oral findings in male prisoners: A systematic review. *J Clin Med*. 2024;13(6):1736. doi:10.3390/jcm13061736
11. Swinkels HM, Jilani TN, Tobin EH. Tuberculosis prevention, control, and elimination. In: *StatPearls*. Treasure Island, USA: StatPearls Publishing; 2025:Bookshelf ID: NBK513246. http://www.ncbi.nlm.nih.gov/books/NBK513246. Accessed September 10, 2025.
12. European Centre for Disease Prevention and Control (ECDC). Tuberculosis: Annual Epidemiological Report for 2022. Solna, Sweden: European Centre for Disease Prevention and Control (ECDC); 2024. https://www.ecdc.europa.eu/en/publications-data/tuberculosis-annual-epidemiological-report-2022. Accessed June 30, 2025.
13. World Health Organization (WHO). Creating supportive conditions to reduce infectious diseases in prison populations. Geneva, Switzerland: World Health Organization (WHO); 2023. https://www.who.int/europe/publications/i/item/WHO-EURO-2023-8182-47950-70944. Accessed June 30, 2025.
14. Penal Reform International (PRI). Global Prison Trends 2022. London, UK: Penal Reform International (PRI); 2023. https://www.penalreform.org/global-prison-trends-2022. Accessed June 30, 2025.
15. Sunderkötter C, Wohlrab J, Hamm H. Scabies: Epidemiology, diagnosis, and treatment. *Dtsch Arztebl Int*. 2021;118(41):695–704. doi:10.3238/arztebl.m2021.0296
16. Delie AM, Bogale EK, Anagaw TF, et al. Global prevalence and predictors of scabies among prisoners: Systematic review and meta-analysis. *BMC Public Health*. 2024;24(1):1894. doi:10.1186/s12889-024-19401-0
17. World Health Organization (WHO). *WHO Informal Consultation on a Framework for Scabies Control, World Health Organization Regional Office for the Western Pacific, Manila, Philippines, 19–21 February 2019: Meeting Report*. Geneva, Switzerland: World Health Organization (WHO); 2020. ISBN:978-92-4-000806-9.
18. Lamichhane B, Mawad AMM, Saleh M, et al. Salmonellosis: An overview of epidemiology, pathogenesis, and innovative approaches to mitigate the antimicrobial resistant infections. *Antibiotics (Basel)*. 2024;13(1):76. doi:10.3390/antibiotics13010076
19. European Centre for Disease Prevention and Control (ECDC). Salmonellosis: Annual Epidemiological Report for 2022. Solna, Sweden: European Centre for Disease Prevention and Control (ECDC); 2024. https://www.ecdc.europa.eu/en/publications-data/salmonellosis-annual-epidemiological-report-2022. Accessed June 30, 2025.
20. Aslam A, Hashmi MF, Okafor CN. Shigellosis. In: *StatPearls*. Treasure Island, USA: StatPearls Publishing; 2025:Bookshelf ID: NBK482337. http://www.ncbi.nlm.nih.gov/books/NBK482337. Accessed September 10, 2025.
21. European Centre for Disease Prevention and Control (ECDC). Shigellosis: Annual Epidemiological Report for 2022. Solna, Sweden: European Centre for Disease Prevention and Control (ECDC); 2024. https://www.ecdc.europa.eu/en/publications-data/shigellosis-annual-epidemiological-report-2022. Accessed June 30, 2025.
22. Government of Poland. The Act of 29 June 1995 on Official Statistics. Warsaw, Poland: Government of Poland; 1995. https://bip.stat.gov.pl/en/law/law-on-public-statistics.
23. Mehta P, Zimba O, Gasparyan AY, Seiil B, Yessirkepov M. Ethics committees: Structure, roles, and issues. *J Korean Med Sci*. 2023;38(25):e198. doi:10.3346/jkms.2023.38.e198
24. Makowski D, Lüdecke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B. Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption. Comprehensive R Archive Network (CRAN); 2023. https://cloud.r-project.org/web/packages/report/index.html. Accessed October 10, 2024.
25. Patil I. Visualizations with statistical details: The "ggstatsplot" approach. *J Open Source Soft*. 2021;6(61):3167. doi:10.21105/joss.03167
26. Wickham H. *Ggplot2*. Cham, Switzerland: Springer International Publishing; 2016. doi:10.1007/978-3-319-24277-4
27. Lambert LA, Armstrong LR, Lobato MN, Ho C, France AM, Haddad MB. Tuberculosis in jails and prisons: United States, 2002–2013. *Am J Public Health*. 2016;106(12):2231–2237. doi:10.2105/AJPH.2016.303423
28. Séri B, Koffi A, Danel C, et al. Prevalence of pulmonary tuberculosis among prison inmates: A cross-sectional survey at the Correctional and Detention Facility of Abidjan, Côte d'Ivoire. *PLoS One*. 2017;12(7):e0181995. doi:10.1371/journal.pone.0181995
29. Ilievska-Poposka B, Zakoska M, Pilovska-Spasovska K, Simonovska L, Mitreski V. Tuberculosis in the prisons in the Republic of Macedonia, 2008–2017. *Open Access Maced J Med Sci*. 2018;6(7):1300–1304. PMID:30087741. PMCID:PMC606227.
30. Seyedalinaghi S, Farhoudi B, Najafi Z, Jafari S, Shahbazi M. Comparing tuberculosis incidence in a prison with the society, Tehran, Iran. *Arch Clin Infect Dis*. 2018;13(2):e60247. doi:10.5812/archcid.60247
31. Jordan AM, Podewils LJ, Castro KG, Zishiri V, Charalambous S. Prevalence and risk factors of tuberculosis disease in South African correctional facilities in 2015. *Int J Tuberc Lung Dis*. 2019;23(11):1198–1204. doi:10.5588/ijtld.18.0782
32. Tong Y, Jiang S, Guan X, et al. Epidemic situation of tuberculosis in prisons in the Central Region of China. *Am J Trop Med Hyg*. 2019;101(3):510–512. doi:10.4269/ajtmh.18-0987
33. Bartosik K, Tytuła A, Zając Z, et al. Scabies and pediculosis in penitentiary institutions in Poland: A study of ectoparasitoses in confinement conditions. *Int J Environ Res Public Health*. 2020;17(17):6086. doi:10.3390/ijerph17176086
34. Amoako YA, Oppong MN, Laryea DO, et al. Burden of scabies in a Ghanaian penitentiary. *PLoS One*. 2024;19(12):e0312108. doi:10.1371/journal.pone.0312108
35. Rogalska A, Barański K, Rachwaniec-Szczecińska Ż, Holecki T, Bąk-Sosnowska M. Assessment of satisfaction with health services among prisoners: Descriptive study. *Healthcare (Basel)*. 2022;10(3):548. doi:10.3390/healthcare10030548
36. Government of Poland. The Act of May 10, 2018, on the Protection of Personal Data (implementation of the GDPR). Warsaw, Poland: Government of Poland; 2018. https://uodo.gov.pl/en/660/1464.
37. World Prisoner Brief (WPB). Highest to Lowest: Occupancy level (based on official capacity). London, UK: Institute for Crime & Justice Policy Research Faculty of Humanities and Social Sciences Birkbeck, University of London; 2024. https://www.prisonstudies.org/highest-to-lowest/occupancy-level?field_region_taxonomy_tid=All. Accessed June 30, 2025.
38. Marco A, García-Guerrero J. Prison overcrowding and over-occupation: What we are talking about and the situation in Spanish prisons. *Rev Esp Sanid Penit*. 2020;22(3):93–95. doi:10.18176/resp.00017
39. Velen K, Charalambous S. Tuberculosis in prisons: An unintended sentence? *Lancet Public Health*. 2021;6(5):e263–e264. doi:10.1016/S2468-2667(21)00049-9
40. Beaudry G, Zhong S, Whiting D, Javid B, Frater J, Fazel S. Managing outbreaks of highly contagious diseases in prisons: A systematic review. *BMJ Glob Health*. 2020;5(11):e003201. doi:10.1136/bmjgh-2020-003201
41. Kendig NE, Bur S, Zaslavsky J. Infection prevention and control in correctional settings. *Emerg Infect Dis*. 2024;30(13):S88–S93. doi:10.3201/eid3013.230705
42. Dolan K, Wirtz AL, Moazen B, et al. Global burden of HIV, viral hepatitis, and tuberculosis in prisoners and detainees. *Lancet*. 2016;388(10049):1089–1102. doi:10.1016/S0140-6736(16)30466-4
43. Korycinska J, Dzika E, Kloch M. Epidemiology of scabies in relation to socio-economic and selected climatic factors in north-east Poland. *Ann Agric Environ Med*. 2020;27(3):374–378. doi:10.26444/aaem/109319
44. Buczek A, Pabis B, Bartosik K, Stanislawek IM, Salata M, Pabis A. Epidemiological study of scabies in different environmental conditions in central Poland. *Ann Epidemiol*. 2006;16(6):423–428. doi:10.1016/j.annepidem.2005.06.058
45. Delaš Aždajić M, Bešlić I, Gašić A, Ferara N, Pedić L, Lugović-Mihić L. Increased scabies incidence at the beginning of the 21st century: What do eeports from Europe and the world show? *Life*. 2022;12(10):1598. doi:10.3390/life12101598
46. Schneider S, Wu J, Tizek L, Ziehfreund S, Zink A. Prevalence of scabies worldwide: An updated systematic literature review in 2022. *Acad Dermatol Venereol*. 2023;37(9):1749–1757. doi:10.1111/jdv.19167
47. Insfran-Rivarola A, Tlapa D, Limon-Romero J, et al. A systematic review and meta-analysis of the effects of food safety and hygiene training on food handlers. *Foods*. 2020;9(9):1169. doi:10.3390/foods9091169
48. Dadu A, Ciobanu A, Hovhannesyan A, et al. Tuberculosis notification trends and treatment outcomes in penitentiary and civilian health care sectors in the WHO European Region. *Int J Environ Res Public Health*. 2021;18(18):9566. doi:10.3390/ijerph18189566